# Public Report - Módulo 2

Aluno: Vinícius Oliveira Fernandes

Orientadora: Cristina Gramani

Este módulo do projeto tem como objetivo aprofundar a análise sobre a imprevisibilidade das recessões econômicas a partir de dados históricos, e o desenvolvimento do modelo de Simulação de Montecarlo.

O trabalho foi estruturado em cinco sprints:

- 1. Coleta e Análise exploratória dos indicadores econômicos
- Primeira Implementação da Simulação de Monte Carlo;
- 3. Refinamento dos parâmetros e cenários da simulação
- 4. Inferência dos resultados obtidos;
- 5. Documentação dos insights gerados.

Neste contexto o desenvolvimento do modelo e suas premissas foram o principal objetivo do módulo, no qual o foco central foi investigar se esses indicadores possuem capacidade preditiva diante de choques econômicos, contribuindo para uma leitura mais probabilística dos ciclos recessivos.

### **Entendimento dos Dados**

- CLI
  - o Dados de 2006 (2006-03-01) até 2025 (2025-03-01)
  - Especificações Estatísticas
    - Média: 5357.988024

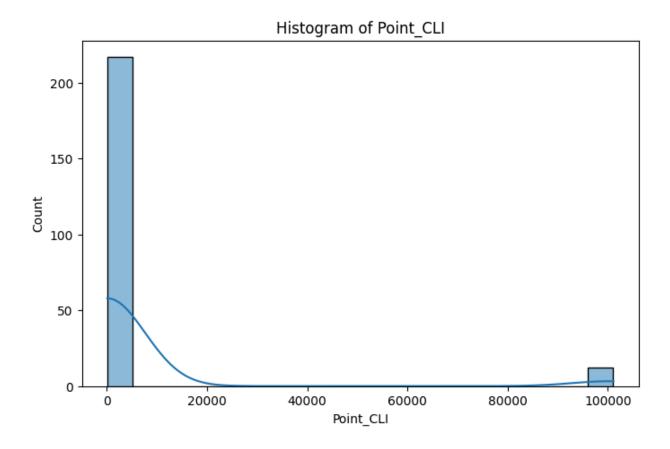
Desvio Padrão: 22409.016195

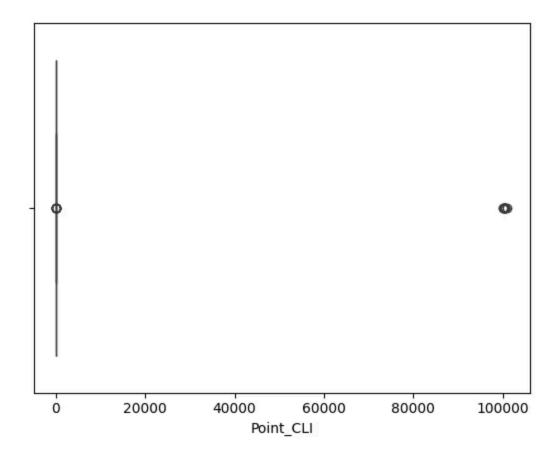
Valor Mínimo: 93.596980

Quartis:

25% → 99.465290

• 50% →100.066300 75% →100.673800





- A maioria dos valores do CLI está concentrada em uma faixa muito estreita (valores baixos), sugerindo um comportamento normal.
  - Há um ponto extremamente discrepante, com valor próximo de 100.000, o que é visivelmente um outlier severo.
    - Isso provavelmente representa erro de entrada ou um dado que foi corrigido/tratado antes da análise estatística.

# • LEI Federal Reserve (USSLIND)

o Dados de 1982 (1982-01-01) até 2020 (2020-02-01)

Especificações Estatísticas

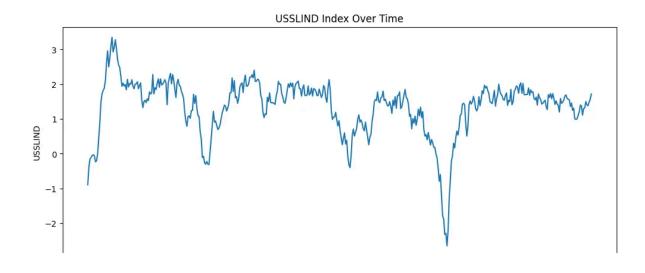
Média: 1.348974

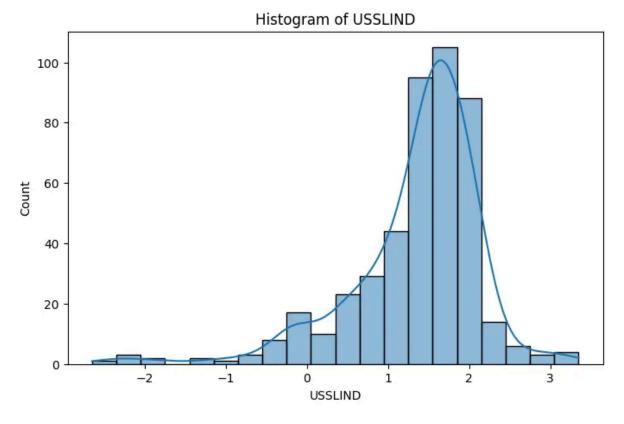
Desvio Padrão: 0.808302

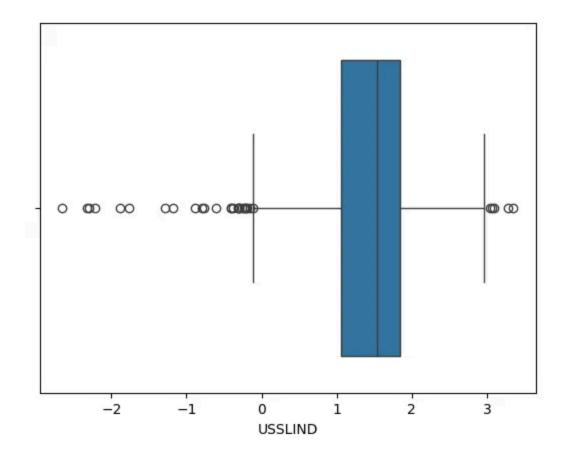
Valor Mínimo: -2.650000

### Quartis:

- 25% → 1.062500
- 50% → 1.535000
- 75% → 1.847500







- A distribuição é assimétrica à esquerda (cauda negativa), mas com um pico concentrado entre 1 e 2, o que indica valores frequentemente positivos.
  - Pequena quantidade de valores negativos, considerados fora do padrão.
  - Vários outliers visíveis, especialmente do lado esquerdo (valores negativos).
  - Mediana está próxima de 1,5.
- Outliers negativos são esperados em momentos de crise ou recessão.

# PIB per capita (EUA)

- o Dados de 1960 até 2023
- Especificações Estatísticas

Média: 29136.216664

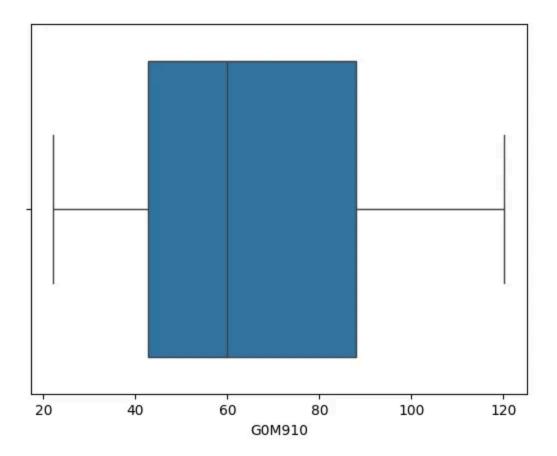
Desvio Padrão: 22026.012318

Valor Mínimo: 2999.864872

- Quartis:
  - 25% → 8394.554319
  - 50% → 24880.624841
  - 75% → 47408.769420

#### • LEI - Conference Board (G0M910)

- Dados mensais de 795 observações
- Especificações Estatísticas
  - Média: 66.288302
  - Desvio Padrão: 26.935913
  - Valor Mínimo: 22.200000
  - Quartis:
    - 25% → 42.800000
    - 50% → 60.000000
    - 75% → 88.050000



- Distribuição simétrica, sem outliers visíveis.
- Valores variam entre 20 e 120, com mediana próxima de 65.
- Essa é uma variável bem comportada do ponto de vista estatístico e pode ser usada diretamente em modelos preditivos sem transformação.

#### • LEI - Conference Board (PCM910)

- Dados mensais de 794 observações
- Especificações Estatísticas

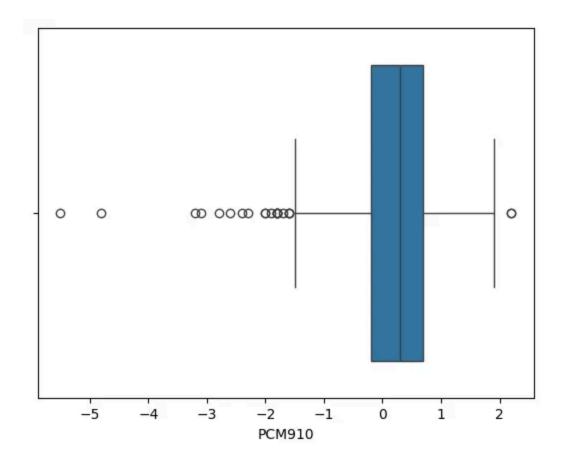
Média: 0.192443

■ Desvio Padrão: 0.748630

Valor Mínimo: -5.500000

#### Quartis:

- 25% → -0.200000
- 50% → 0.300000
- 75% → 0.700000



- Distribuição assimétrica com vários outliers negativos (até -5%).
- A mediana está próxima de 0, sugerindo que em média o índice varia pouco mês a mês.
- Os outliers podem corresponder a eventos econômicos abruptos (choques de mercado, pandemias, crises).

# Inferências e Transformação

| Indicador | Outliers?           | Correção<br>Necessária?    | Notas Importantes                                  |
|-----------|---------------------|----------------------------|--|
| Point_CLI | Sim (bem relevante) | Sim (normalização)         | Outlier distorce completamente a escala            |
| USSLIND   | Sim                 | Não, se<br>contextualizado | Outliers refletem crises econômicas                |
| PCM910    | Sim (pouco)         | Avaliar contexto           | Variações percentuais fortes em choques econômicos |
| G0M910    | Não                 | Não                        | Estável e simétrica                                |

A partir da concatenação em um só dataframe destes dados houve a normalização dos dados de utilizando o MinMaxScaler (X' = (X - X\_min) / (X\_max - X\_min)). O único indicador que não passou pela normalização após testes, foi o PIB. A normalização estava bagunçando muito as simulações que foram rodadas.

### 2008

Período de 24 meses

#### CLI

| Data                | CLI      |
|---------------------|----------|
| 2007-06-01 00:00:00 | 101.7681 |
| 2007-07-01 00:00:00 | 101.7095 |
| 2007-08-01 00:00:00 | 101.6182 |
| 2007-09-01 00:00:00 | 101.5185 |
| 2007-10-01 00:00:00 | 101.4225 |
| 2007-11-01 00:00:00 | 101.3274 |
| 2007-12-01 00:00:00 | 101.2353 |

| 2008-01-01 00:00:00  | 101.1377  |
|--|---|
| 2008-02-01 00:00:00  | 101.0243  |
| 2008-03-01 00:00:00  | 100.8861  |
| 2008-04-01 00:00:00  | 100.7035  |
| 2008-05-01 00:00:00  | 100448  |
| 2008-06-01 00:00:00  | 100.0939  |
| 2008-07-01 00:00:00  | 99.62764  |
| 2008-08-0100:00:00   | 99.04966  |
| 2008-09-0100:00:00   | 98.37462  |
|  |   |
| 2008-10-01 00:00:00  | 97.64918  |
| 2008-10-01 00:00:00<br>2008-11-01 00:00:00   | 97.64918<br>96.9623   |
|  | 07.0 10.10  |
| 2008-11-01 00:00:00  | 96.9623   |
| 2008-11-01 00:00:00 2008-12-01 00:00:00  | 96.9623<br>96.39452   |
| 2008-11-01 00:00:00<br>2008-12-01 00:00:00<br>2009-01-01 00:00:00  | 96.9623<br>96.39452<br>96.00105                                     |
| 2008-11-01 00:00:00<br>2008-12-01 00:00:00<br>2009-01-01 00:00:00<br>2009-02-01 00:00:00   | 96.9623<br>96.39452<br>96.00105<br>95.81349                         |
| 2008-11-01 00:00:00<br>2008-12-01 00:00:00<br>2009-01-01 00:00:00<br>2009-02-01 00:00:00<br>2009-03-01 00:00:00                        | 96.9623<br>96.39452<br>96.00105<br>95.81349<br>95.84404             |
| 2008-11-01 00:00:00<br>2008-12-01 00:00:00<br>2009-01-01 00:00:00<br>2009-02-01 00:00:00<br>2009-03-01 00:00:00<br>2009-04-01 00:00:00 | 96.9623<br>96.39452<br>96.00105<br>95.81349<br>95.84404<br>96.07906 |

Os únicos "ofensores" que estariam abaixo do primeiro quartil (25%) contendo os valores históricos mais baixos são os dados do começo de 2009 até o meio do mesmo ano (aproximadamente 6 meses). Este período pode ser considerado como o "pós fato" da recessão em si, nos quais a economia já realmente está sofrendo com o evento de recessão em si.

# **LEI - Conference Board**

| Date    | G0M910 | PCM910 | P6M910 |
|---------|--------|--------|--------|
| 2007-06 | 97,1   | -0,2   | -3,5   |
| 2007-07 | 96,9   | -0,2   | -3,7   |

| 2007-08 | 96,5 | -0,4 | -5,3  |
|---------|------|------|-------|
| 2007-09 | 95,9 | -0,6 | -6,7  |
| 2007-10 | 95,5 | -0,4 | -8,5  |
| 2007-11 | 94,7 | -0,8 | -9,9  |
| 2007-12 | 93,8 | -1,0 | -11,3 |
| 2008-01 | 92,7 | -1,2 | -11,6 |
| 2008-02 | 91,6 | -1,2 | -11,7 |
| 2008-03 | 90,3 | -1,4 | -10,8 |
| 2008-04 | 89,8 | -0,6 | -12,7 |
| 2008-05 | 89,0 | -0,9 | -12,7 |
| 2008-06 | 88,6 | -0,4 | -13,7 |
| 2008-07 | 86,6 | -2,3 | -18,2 |
| 2008-08 | 85,6 | -1,2 | -21,0 |
| 2008-09 | 83,9 | -2,0 | -24,7 |
| 2008-10 | 81,2 | -3,2 | -24,2 |
| 2008-11 | 79,1 | -2,6 | -24,5 |
| 2008-12 | 76,9 | -2,8 | -23,9 |
| 2009-01 | 75,4 | -2,0 | -18,5 |
| 2009-02 | 74,4 | -1,3 | -13,2 |
| 2009-03 | 73,2 | -1,6 | -6,6  |
| 2009-04 | 73,3 | 0,1  | -1,8  |
| 2009-05 | 73,7 | 0,5  | 2,7   |
| 2009-06 | 74,3 | 0,8  | 8,1   |

# Explicação rápida de cada específicação dentro do LEI:

| G0M910 | Nível do índice LEI no mês atual                     | Estado atual do índice             |
|--------|--|------------------------------------|
| PCM910 | Variação percentual mensal do LEI                    | Mudança econômica mês a mês        |
| P6M910 | Variação percentual acumulada nos<br>últimos 6 meses | Tendência econômica no médio prazo |

O LEI já é perceptivamente mais sensível do que os outros indicadores, tendo em vista sua adaptabilidade em relação à variações abruptas. Neste sentido os indicadores de variação: PCM910 (variação mensal) e P6M910 (variação de 6 meses), acabam se destacando na identificação de flutuações "anomálicas" e neste caso é possível perceber que os primeiros indícios de desaceleração econômica, e possibilidade de bolha, perto do final de 2007, já começam a ser identificados, mês de dezembro por exemplo:

| Date    | G0M910 | PCM910 | P6M910 |
|---------|--------|--------|--------|
| 2007-12 | 93,8   | -1,0   | -11,3  |

#### **US LEI Federal Reserve Bank**

| 2007-06-01 | 0.70  |
|------------|-------|
| 2007-07-01 | 0.51  |
| 2007-08-01 | 0.55  |
| 2007-09-01 | 0.40  |
| 2007-10-01 | 0.61  |
| 2007-11-01 | 0.43  |
| 2007-12-01 | 0.25  |
| 2008-01-01 | 0.41  |
| 2008-02-01 | 0.32  |
| 2008-03-01 | 0.19  |
| 2008-04-01 | 0.18  |
| 2008-05-01 | 0.00  |
| 2008-06-01 | -0.12 |
| 2008-07-01 | -0.41 |
| 2008-08-01 | -0.79 |
| 2008-09-01 | -0.60 |
| 2008-10-01 | -1.18 |

| 2008-11-01 | -1.77 |
|------------|-------|
| 2008-12-01 | -1.88 |
| 2009-01-01 | -2.32 |
| 2009-02-01 | -2.29 |
| 2009-03-01 | -2.65 |
| 2009-04-01 | -2.21 |
| 2009-05-01 | -1.29 |
| 2009-06-01 | -0.77 |

USLIND, tmabém é uma extensão do LEI, mas ele serve para prever como a economia de um estado vai crescer nos próximos seis meses. Ele faz isso olhando para um outro índice chamado "índice coincidente", que mostra a situação atual da economia do estado. Ou seja, o USLIND tenta antecipar se a economia vai melhorar ou piorar no curto prazo.

Os dados do USLIND indicam uma desaceleração econômica no começo de 2008. Esse indicativo pode ser considerado pré-recessão pois o "boom" de recessão no período foi um pouco mais tarde. Em cunho previsivo, foi até que relevante dentro do período em si.

### 2020

Período de 12 meses

#### CLI

| 2020-01-01 00:00:00 | 99.1925  |
|---------------------|----------|
| 2020-02-01 00:00:00 | 99.17662 |
| 2020-03-0100:00:00  | 97.49301 |
| 2020-04-0100:00:00  | 93.59698 |
|                     |          |
| 2020-05-01 00:00:00 | 95.07487 |

| 2020-07-01 00:00:00 | 98.29512 |
|---------------------|----------|
| 2020-08-01 00:00:00 | 98.77375 |
| 2020-09-01 00:00:00 | 98.97484 |
| 2020-10-01 00:00:00 | 99.21239 |
| 2020-11-01 00:00:00 | 99.47532 |
| 2020-12-01 00:00:00 | 99.8277  |
| 2021-01-01 00:00:00 | 100.0862 |

O CLI apenas "sentiu" o período de maior incidência do impacto e da surpresa da pandemia. No mês de março, na qual foi identificado o nível alto de incidência e casos pelo mundo. Depois em Abril, o caos tomou conta do mercado (é possível notar a baixa de 93.5 pontos), a partir da surpresa do impacto da doença. Até o final se setembro, o indíce se manteve abaixo do primeiro quartil histórico, determinando um período de "recessão".

### **LEI (Conference Board):**

| Date    | G0M910 | PCM910 | P6M910 |
|---------|--------|--------|--------|
| 2020-02 | 112,9  | 0,0    | -15,8  |
| 2020-03 | 107,5  | -4,8   | -11,6  |
| 2020-04 | 101,6  | -5,5   | -9,3   |
| 2020-05 | 103,5  | 1,9    | -7,6   |
| 2020-06 | 105,8  | 2,2    | 3,0    |
| 2020-07 | 107,5  | 1,6    | 16,6   |
| 2020-08 | 108,5  | 0,9    | 13,8   |
| 2020-09 | 109,1  | 0,6    | 10,3   |
| 2020-10 | 109,7  | 0,5    | 8,5    |
| 2020-11 | 110,4  | 0,6    | 6,6    |
| 2020-12 | 111,1  | 0,6    | 7,5    |
| 2021-01 | 112,0  | 0,8    | 8,6    |
| 2021-02 | 112,0  | 0,0    | 8,7    |

| 2021-03 | 113,1 | 1,0 | 9,2 |
|---------|-------|-----|-----|
|         | '     | •   | '   |

A mesma tendência é perceptível no LEI, na qual o período de março até setembro é ainda um período ofensivo, em termos econômicos, para o índice geral (*G0M910*), no entantoas flutações mensais (*PCM910*) e as de 6 em seis meses (*P6M910*) refletem a uma variação abrupta alguns meses antes da definição de estado pandêmicos, e também nos meses de maior incerteza e surpresa econômica com a chegada do Vírus (Abril, Março e Maio), também no período exato de maior incerteza econômica.

# Premissa Definida para as Simulações

#### Definição do que poderia se enquadrar em uma recessão:

- Alto Percentual de queda nos indicadores financeiros definidos como inputs do modelo
  - Análise dos quartis e seleção dos 0.25 mínimos de cada indicador como identificação de uma mudança na tendência
- "Flag" definida a partir de premissa na análise dos dados históricos de cada indicador, assim como análise do comportamento desses indicadores dentro de uma recessão.

Atribuição da Simulação de Monte Carlo: Simular possíveis trajetórias futuras para estes indicadores baseado em padrões históricos.

- Execução das Simulações de Monte Carlo → Para cada situação, nesse caso de países, simulado, verificar se ocorre uma recessão.
- "Probabilidade" de uma recessão → Com base nos resultados das simulações, é feito o calculo da probabilidade de uma recessão.

### Pré simulação

#### **Definição dos limiares (thresholds)**

- Definição dos limiares (thresholds) de cada indicador com base no 1º quartil
   (25%) da sua distribuição histórica.
- Esses valores servem como referência para detectar níveis economicamente
   "baixos" que indicam uma recessão.

Obs: Apesar desses thresholds estarem calculados, a função is\_recession usa valores fixos (cli < 100, g0 < 100, pcm < 0.25) ao invés dos thresholds; Isso foi feito por ter evidênciado resultados ruins nas rodadas de simulação

Função de definição dos marcadores de recessão:

```
def is_recession(cli, g0, pcm):
    return (
        cli < limiar_025 and
        g0 < limiar_025 and
        pcm < limiar_025
)</pre>
```

#### Marcadores mensais

```
returns = df_merged[['Point_CLI', 'G0M910', 'PCM910']].pct_change().dropna()
```

- A intenção é pegar a variação percentual mensal (retorno) dos três indicadores econômicos.
  - Isso representa o comportamento histórico de cada variável ao longo do tempo.
  - A distribuição desses retornos é usada para simular o comportamento futuro.

# Na simulação

Função da simulação:

def monte\_carlo\_simulation(current\_values, num\_months, num\_simulations=1000

#### Motorzinho da simulação:

- Cada simulação começa com os valores atuais dos indicadores ( current\_values ).
- Inicializa a flag recession\_occurred como False.

Mais um nível de aleatoriedade adicionado, um **choque** aleatório é aplicado em cada mês:

- Um vetor de retornos ( shock ) é sorteado aleatoriamente da série histórica ( returns ).
- Esse vetor contém variações percentuais para Point\_CLI, GOM910 e PCM910.
- Os valores dos indicadores são atualizados multiplicando pelo fator de choque:novo valor=valor atual×(1+retorno simulado)
   novo valor=valor atual×(1+retorno simulado)\text{novo valor} = \text{valor atual} \times (1 + \text{retorno simulado})
- Isso simula o comportamento futuro dos indicadores com base nos padrões históricos de volatilidade.

#### Pós Fato:

- Após a aplicação de cada choque mensal, o código verifica se os indicadores caíram abaixo dos limiares de recessão.
- Se isso ocorrer em qualquer mês da simulação, o flag recession\_occurred é ativado.

- Cada simulação termina com True (se houve recessão) ou False (caso contrário).
- Ao final, o retorno é uma lista com 100.000 valores booleanos.
- Como True é tratado como 1 e False como 0, a média da lista retorna a proporção de simulações que indicaram recessão.
- Isso representa a probabilidade estimada de uma recessão ocorrer nos próximos n meses.

### Reporte dos Resultados

#### United States - 2008 (Crise do SubPrime)

Utilizando os primeiros quartis de cada um dos indicadores como threshold para definição de recessão ou não (baseado na identificação das tendências e variações dos dados históricos):

- Probababilidade estimada de 12 meses, 61% de indício de recessão:
  - Dados históricos 12 meses → 2007-06-12 até 2008-06-12

 Utilizando dados históricos apenas do ano de 2007 → Probabilidade de recessão em 6 meses: 89,32%

Dados históricos 12 meses → 2007 inteiro

A partir de algumas rodades e investigação, foi identificado que ao utilizar os quartis de uma amostra menor, os resultados são melhores.

Como "bônus" rodei também uma bateria de testes com um limiar amplo para a recessão de 2020:

• Resultado de 9% de indício de recessã

```
United States - 2020

| Comment of the profit are post that 1,00 control (1920 to 1920 to 1920
```

### **United States – 2020 (Pandemia)**

- · Janela de 6 meses:
  - o 2019-12 | 2020-01 | 2020-02 | 2020-03 | 2020-04 | 2020-05

```
current_values_2020 = us_2020[['Point_CLI', 'GOM910', 'PCM910']].iloc[-1].values

returns_2020 = monte_carlo_simulation(current_values_2020)
    prob_recession_2020 = np.mean(returns_2020)

print(f"Estimated probability of recession in next months: {prob_recession_2020:.2%}")

very 2m 9.3s

/tmp/ipykernel_28238/4168068882.py:8: RuntimeWarning: invalid value encountered in multiply values *= (1 + shock)
Estimated probability of recession in next months: 20.68%
```

- Janela de 9 Meses:
  - 2019-09 | 2019-10 | 2019-11 | 2019-12 | 2020-01 | 2020-02 | 2020-03 | 2020-05

```
current_values_2020 = us_2020[['Point_CLI', 'GOM910', 'PCM910']].iloc[-1].values
returns_2020 = monte_carlo_simulation(current_values_2020)
prob_recession_2020 = np.mean(returns_2020)

print(f"Estimated probability of recession in next months: {prob_recession_2020:.2%}")

/ 2m 11.6s

/tmp/ipykernel_28238/4168068882.py:8: RuntimeWarning: invalid value encountered in multiply values *= (1 + shock)
Estimated probability of recession in next months: 3.21%
```

#### • Janela de 12 Meses:

2019-07 | 2019-08 | 2019-09 | 2019-10 | 2019-11 | 2019-12 | 2020-01 |
 2020-02 | 2020-03 | 2020-04 | 2020-05 | 2020-06

#### Janela de 24 Meses:

2019 and 2020

```
current_values_2020 = us_2020[['Point_CLI', 'G0M910', 'PCM910']].iloc[-1].values

returns_2020 = monte_carlo_simulation(current_values_2020)
prob_recession_2020 = np.mean(returns_2020)

print(f"Estimated probability of recession in next months: {prob_recession_2020:.2%}")

/ 2m 12.4s

/tmp/ipykernel_28238/4168068882.py:8: RuntimeWarning: invalid value encountered in multiply values *= (1 + shock)
Estimated probability of recession in next months: 0.07%
```