

# 06

## **Avaliação dos Incentivos Fiscais da Sudene como instrumento promotor do desenvolvimento regional**

**Produto 6 - Glossário com os termos das avaliações**

## **Expediente**

### **Presidência da República Federativa do Brasil**

Luiz Inácio Lula da Silva

### **Ministério da Integração e Desenvolvimento Regional**

Antônio Waldez Góes da Silva

### **Superintendência do Desenvolvimento do Nordeste**

Danilo Jorge de Barros Cabral

### **Diretoria de Planejamento e Articulação de Políticas**

Álvaro Silva Ribeiro

### **Coordenação Geral de Estudos e Pesquisas, Avaliação, Tecnologia e Inovação**

Rafaella Iliana Alves Arcila

### **Coordenação de Avaliação e Estudos**

Maria da Glória Carré Martins Sistêlos

### **Equipe Técnica**

Diana Dias Sampaio

Gabriela Isabel Limoeiro Alves Nascimento

Marina Rogério de Melo Barbosa

Miguel Vieira Araújo

### **Universidade Federal do Ceará**

Custódio Luís Silva de Almeida

Reitor

Diana Cristina Silva de Azevedo

Vice-reitora

### **Departamento de Economia Aplicada**

Guilherme Irffi

Chefe do Departamento

Glauber Nojosa

Subchefe do Departamento

### **Projeto**

Avaliação dos Resultados dos incentivos fiscais sob gestão da Sudene

### **Coordenador Geral do Projeto**

Guilherme Irffi

**Pesquisadores**

Ana Sara Ribeiro Parente Cortez Irffi  
Diego Rafael Fonseca Carneiro  
Edward Martins Costa  
Felipe de Souza Bastos  
Georgeana Amaral Maciel da Silveira  
Maitê Rimekká Shirasu

**Assistentes de Pesquisa**

Ana Caroline de Freitas Peixoto  
Ana Karine Justino da Costa  
Antônio Lucas de Abreu Melo  
Brysa Fernandes  
Camila Guedes Corrêa  
Erivelton Nunes  
Gabriella Costa  
Isabela Braga Sales  
Jean Haendell  
José Carlos Souza Lima  
Juliana Cavalcanti  
Maria Vanessa Andrade Silva  
Marleton Souza Braz  
Kelly Samá Lopes de Vasconcelos  
Rayane Costa Dantas  
Pedro Alexandre Veloso  
Sarah Alves  
Thyena Karen Magalhães Dias  
Wendel Mendes Ferreira

**Universidade Federal do Ceará - UFC**

**Superintendência do Desenvolvimento do Nordeste - SUDENE**

# Glossário de Avaliação Econômica

## A

**Ação:** São os processos que, combinando apropriadamente os recursos adequados, produzem bens e serviços com os quais se procura atacar as causas do problema.

**Análise Custo-Benefício:** Termo usado para denominar uma diversidade de técnicas que buscam fornecer parâmetros para estudar a viabilidade econômica de um projeto e a conveniência de sua execução, em comparação com as demais alternativas possíveis. O programa é considerado economicamente viável se os benefícios totais excederem seus custos totais.

**Análise Custo Efetividade:** É uma alternativa para relacionar o impacto gerado e seus custos quando não é possível/pertinente calcular o benefício de um projeto (quando a monetização do impacto de um projeto social pode ser muito complexa). A Análise Custo Efetividade apenas compara a eficiência de diferentes intervenções (que tenham o mesmo indicador de impacto).

**Análise de Sensibilidade:** Processo de verificar se os efeitos estimados e a significância estatística de variáveis explicativas cruciais são sensíveis à inclusão de outras variáveis explicativas, à forma funcional, à eliminação de observações potencialmente atípicas, ou a diferentes métodos de estimação.

**Atrito:** Ocorre quando na coleta de dados de campo em mais de um momento do tempo há perda de informações sobre alguns indivíduos que não são encontrados nos períodos subsequentes ao primeiro (atrito amostral).

**Avaliação de Impacto:** Verifica se o programa está alcançando os objetivos ao qual se propôs. O impacto é o contraste entre duas situações, uma real (situação dos beneficiários do projeto após sua participação) e outra hipotética (situação dos participantes caso não tivessem participado do programa). Dividida em duas categorias, método experimental e método não experimental.

**Avaliação de processos:** Avaliação da dinâmica interna dos organismos responsáveis pela implementação da intervenção, das suas

políticas, das suas práticas de gestão e da ligação entre todas estas componentes.

**Avaliação econômica:** Consiste em quantificar o impacto e o retorno econômico de um projeto, com base em metodologia científica.

**Avaliação Ex-Ante:** Trata-se de uma avaliação realizada antes da ocorrência da ação; isto é, deve ser realizada com base na construção de um modelo sobre o comportamento dos agentes e do ambiente econômico no qual eles e o programa em avaliação estão inseridos, geralmente é empregado antes da implementação da intervenção. Em alguns casos, o modelo proposto é estrutural, ou seja, baseado em hipóteses acerca das formas funcionais que representam o comportamento econômico dos indivíduos e do funcionamento da economia ao longo do tempo. Frequentemente, essa avaliação utiliza dados de fontes secundárias ao programa para estimar os parâmetros do modelo proposto, incluindo o próprio impacto do programa.

**Avaliação Ex-Post:** Esse tipo de avaliação pode ser realizado durante ou após a execução do programa. E, para isto, utilizam-se os dados da intervenção para explicar os resultados encontrados.

**Avaliação externa:** Metodologia que permite generalizar as estimativas do impacto do programa avaliado para outras populações e outros momentos no tempo. Uma forma de garantir a validade externa dos resultados em certo ponto no tempo é que a amostra utilizada na avaliação seja representativa da população-alvo do programa naquele momento.

**Avaliação interna:** Questiona se o desenho da avaliação identifica corretamente o efeito causal de um programa. Quando um estudo possui validade interna, seus resultados conseguem isolar o verdadeiro impacto do programa, isto é, estão livres da influência de outros fatores que não o próprio efeito da intervenção.

**Avaliação Qualitativa:** Quando se deseja compreender quais os significados que os sujeitos atribuem a qualquer uma das dimensões do programa (definição da demanda, objetivos, estratégias, metas, resultados, sustentabilidade etc.) ou como se dá, na prática e no cotidiano, a interface entre a realização das ações do programa e os discursos/concepções sobre estas ações.

**Avaliação Quantitativa:** Análise que envolve fatores mensuráveis, com uso de dados como percentagens, média, mediana, coeficiente de correlação, entre outros.

# B

**Benchmark:** Expressão em inglês que significa “ponto de referência” ou “unidades-padrão”, para que se estabeleçam comparações entre produtos, serviços, processos, títulos, taxas de juros etc.

**Benefício:** é o valor monetário do impacto causado por um projeto social ou política pública.

**Beneficiário direto:** Indivíduos que participam das atividades desenvolvidas no âmbito do programa.

**Beneficiário indireto:** Indivíduos influenciados pela ação dos beneficiários diretos do programa.

**Bottom-Up:** Procura de resultados extraordinários para ações específicas antes de considerar o impacto das tendências econômicas.

**Brainstorming:** Termo em inglês que significa, literalmente, tempestade ou tormenta cerebral, isto é, um esforço concentrado da inteligência do pessoal mais qualificado na busca da solução de um problema.

# C

**Contrafactual:** um grupo de comparação que seja o mais próximo possível do que teria ocorrido com o grupo tratado caso este não tivesse recebido a intervenção.

**Critérios de elegibilidade:** são os critérios estabelecidos pelo programa para classificar o público-alvo como elegível. Em outros termos, são as características que o público-alvo precisa para se tornar beneficiário do programa.

**Custo-Benefício:** É uma relação que pode ser favorável ou desfavorável. Isto é, quando favorável o valor dos benefícios do programa supera os custos do programa. É um indicador que relaciona os benefícios e os custos do programa em termos monetários.

**Custo Contábil:** é o custo advindo da contabilidade do projeto, tudo que foi efetivamente desembolsado.

**Custo de oportunidade:** é o rendimento que se deixa de obter quando se realiza uma determinada escolha.

**Custo Efetividade:** É um indicador que relaciona o impacto do programa por unidade de custo. Segundo o exemplo citado por Menezes Filho et al. (2012) se em uma sociedade existirem dois programas que possuem o mesmo objetivo a escolha por um determinado programa para atuar na sociedade será aquele que produz o maior impacto por unidade de custo.

## D

**Dados primários:** são informações coletadas a partir de questionários e entrevistas pelo próprio pesquisador

**Dados secundários:** são informações coletadas por dados secundários, coletados por órgãos governamentais como, por exemplo, IBGE, INEP, MTE.

**Diferença-em-diferenças:** método baseado no cálculo de uma dupla subtração, em que a primeira se refere à diferença das médias da variável de resultado entre os períodos anterior e posterior ao programa, para o grupo de tratamento ( $\Delta Y_{T_1, t_1-t_0}$ ) e para o de controle ( $\Delta Y_{T_0, t_1-t_0}$ ), e a segunda se refere a diferença da primeira diferença calculada entre esses dois grupos. Denotando por  $T = \{1,0\}$  a participação ou não no programa e por  $t = \{1,0\}$  os períodos posterior e anterior a intervenção, respectivamente, o estimador de DD será dado por:  $\beta_{DD} = \{E[Y|T = 1, t = 1] - E[Y|T = 1, t = 0]\} - \{E[Y|T = 0, t = 1] - E[Y|T = 0, t = 0]\}$ .

## E

**Efeito causal:** É o impacto de um programa sobre os beneficiários do projeto. A avaliação desse impacto na vida dos beneficiários necessita de indicadores objetivos e mensuráveis que possam ser utilizados para aferir o efeito, ou seja, o sucesso desse programa.

**Efeito direto:** Representa o conjunto de resultados a curto e médio prazo esperados ou alcançados pela intervenção.

**Efeito Hawthorn:** Tratamento pode modificar comportamento do grupo de tratamento.

**Efeito indireto:** Efeito de uma intervenção pública que se propaga na economia, na sociedade ou no ambiente, para além dos seus destinatários diretos.

**Efeito John Henry:** Tratamento pode modificar comportamento do grupo de controle.

**Efeito líquido:** Efeito realmente imputável a uma intervenção pública.

**Efeito transbordamento (Spillover effect):** Também conhecido como contaminação do grupo de comparação. Um efeito de transbordamento ocorre quando o grupo de comparação é afetado pelo tratamento administrado ao grupo de tratamento, embora o tratamento não seja administrado diretamente ao grupo de controle. Se os efeitos de transbordamento sobre o grupo de comparação são negativos (ou seja, o grupo é prejudicado por causa do programa), então a diferença direta entre os resultados dos grupos de tratamento e de comparação gerarão uma superestimação do impacto do programa. Em contrapartida, se o efeito de transbordamento no grupo de comparação é positivo (ou seja, gera benefícios), então ele irá produzir uma subestimação do impacto do programa.

**Efeito:** Alteração intencional ou não intencional devida, direta ou indiretamente, a uma intervenção.

**Efeitos experimentais:** são efeitos provenientes de estudos experimentais que procuram estabelecer uma relação de causa e efeito entre variáveis em estudo de forma prática.

**Efetividade:** é uma medida de duração de um impacto. Isto é, se o projeto social for efetivo, se ele deixar de existir os resultados alcançados não serão desfeitos.

**Eficácia:** é uma medida relacionada a uma meta. Diz-se que um projeto social foi eficaz se ele alcançou um objetivo mensurável por um indicador de impacto e/ou resultado.

**Eficiência:** é uma medida relativa, usada para comparar duas ações de um mesmo projeto ou dois projetos diferentes.

**Eficiência alocativa:** Situação em que os recursos disponíveis para os serviços são utilizados de forma a maximizar os benefícios para a sociedade.

**Eficiência distributiva:** Pode ser definida da mesma forma que eficiência alocativa.

**Ensaio clínico:** uma investigação destinada a descobrir ou verificar os efeitos clínicos, farmacológicos ou os outros efeitos de um ou mais experimentos, ou a analisar a absorção, a distribuição, o metabolismo e a eliminação de um ou mais medicamentos experimentais, a fim de apurar a respectiva segurança ou eficácia.

**Erros de Focalização:** ineficiência no alcance ou na cobertura do programa e vazamento do programa.

**Estudo caso-controle:** quando se usa um grupo de comparação sem a doença ou o desfecho de interesse.

**Estudo transversal:** Um estudo que examina a relação entre doenças (ou outra característica relacionada à saúde) e outras variáveis de interesse sobre o modo como elas se manifestam em uma população definida, em um determinado momento. A sequência temporal de causa e efeito não será, necessariamente, determinada em um estudo transversal.

**Externalidade:** Pode ser positiva ou negativa. Em Políticas Públicas a externalidade pode ser explicada como o transbordamento dos efeitos do programa sobre o grupo de controle. De acordo Menezes Filho et al. (2012), por exemplo, um programa de combate a obesidade infantil, onde os grupos de tratamento e controle foram escolhidos aleatoriamente dentro da mesma escola, pode ocorrer que o tratamento acabe influenciando indiretamente o grupo de controle, visto que as crianças tratadas podem influenciar o comportamento alimentar das crianças do grupo de controle por conviverem no mesmo ambiente.

## F

**Fila de espera:** sequência de pessoas dispostas de maneira alinhada pelos mais diversos critérios (ordem de chegada, altura, risco de morte etc.) e para os mais diversos objetivos.

**Fluxo de Caixa:** É um instrumento de gestão que projeta para períodos futuros todas as entradas e saídas de benefícios ou recursos monetários de um programa ou empresa.

**Focalização:** o programa inclui de fato as pessoas que deveria incluir.

## G

**Gestão orientada para resultados:** Estratégia de gestão que se centra no desempenho, na obtenção de resultados, efeitos e impactos.

**Grupo de Controle:** formado pelos elegíveis que não participam (não recebem o benefício) do programa.

**Grupo de Tratamento:** formado pelos elegíveis que participam (recebem o benefício) do programa.

**Grupo de tratados:** São os beneficiários do programa, ou seja, indivíduos que passaram pela intervenção do projeto.

**Grupo dos não tratados:** Indivíduos que não passaram pela intervenção do programa.

**Grupo Focal:** é uma técnica de entrevista direcionada a um grupo que é organizado a partir de determinadas características identitárias, visando obter informações qualitativas organizadas segundo um determinado quadro teórico de referência.

**Indicador de Impacto:** Reflete a efetividade do projeto em atingir os objetivos propostos.

**Indicador de Produto:** Parte das ações realizadas que se traduzem em algo concreto.

**Indicador de Resultado:** reflete o sucesso das ações realizadas pelo projeto.

**Indicador:** Variável que mede um fenômeno de interesse do avaliador. Pode ser quantitativo ou qualitativo. O indicador informa sobre a evolução empírica de um aspecto observado.

**Ineficiência na cobertura do programa:** consiste em excluir ou não incluir no programa pessoas pertencentes ao público-alvo.

**Insumo (Input):** Os recursos financeiros, humanos e materiais utilizados pelo programa para realizar suas intervenções com os participantes.

L

**LATE:** representa apenas o efeito médio do tratamento para uma subpopulação bem específica, a população de indivíduos que têm seu comportamento alterado em função de uma variação no instrumento. Em geral, este grupo de indivíduos não é representativo da população como um todo.

**Linha de base:** Pré-intervenção, ex-ante. Situação anterior à intervenção, contra a qual o progresso pode ser avaliado ou comparações serem feitas.

## M

**Mapa de processos:** Representação gráfica da sequência de ações que ocorrem entre processos.

**Marco Lógico:** Ferramenta que deve facilitar o processo de justificativa, elaboração, execução e avaliação de projetos em geral (sociais em particular).

**Meta:** Projeção futura de realização em tempo determinado, numa área de reconhecida responsabilidade. Uma meta envolve a conversão de objetivos em tarefas que sejam faseadas no tempo, possíveis, quantificáveis e alcançáveis.

**Meta-análise:** é uma técnica estatística especialmente desenvolvida para integrar os resultados de dois ou mais estudos independentes, sobre uma mesma questão de pesquisa, combinando, em uma medida resumo, os resultados de tais estudos.

**Método experimental:** Seleciona de forma aleatória os grupos de tratados e controles e compara médias dos resultados dos grupos

**Método não experimental:** Substitui a forma aleatória de selecionar os indivíduos tratados e controle do método experimental por certas condições e hipóteses que visam a tornar o grupo de controle parecido com o grupo de tratamento. Os principais métodos não experimentais são: diferenças em diferenças (DD), regressão simples, pareamento e variáveis instrumentais.

**Monetização:** é um exercício de simplificação/ reducionismo que visa transformar/ traduzir algo em dinheiro.

## O

**Objetivos:** significa o fim que se deseja atingir, a meta que se pretende alcançar ou o que é relativo ao objeto, que é concreto e existe independentemente do pensamento. É o que move o indivíduo para tomar alguma decisão.

**Outcome:** objetivos de curto prazo ou intermediários.

**Output:** avaliação que se concentra nos resultados.

# P

**Plano de avaliação:** Documento escrito que antecede a realização de uma avaliação e que poderá incluir as seguintes componentes, entre outras: contexto e modelo lógico de avaliação; finalidade da avaliação; questões de avaliação; métodos de avaliação; equipa de avaliação; procedimentos e logística.

**Plano de monitoramento:** É um processo rotineiro de acúmulo de informações do projeto em todos os seus aspectos para checar o progresso das atividades, servindo como método de fiscalização e instrumento para avaliação.

**Políticas Públicas:** são conjuntos de programas, ações e atividades desenvolvidas pelo Estado diretamente ou indiretamente, com a participação de entes públicos ou privados, que visam assegurar determinado direito de cidadania, de forma difusa ou para determinado segmento social, cultural, étnico ou econômico.

**População:** o grupo completo de observações de pessoas, objetos ou coisas.

**Programa:** Conjunto de projetos e atividades articulados entre si, que permitem alcançar determinados objetivos, consubstanciando a estratégia/política previamente delineada.

**Projetos focalizados:** São dirigidos a um determinado grupo da sociedade, denominado público-alvo.

**Público-alvo:** Composto por indivíduos, empresas, municípios, estados e/ou regiões que o programa (projeto, lei) pretende atender. Sendo assim, o público-alvo consiste em um segmento específico da sociedade com características em comum a qual se destina o programa (projeto e/ou a lei).

# R

**Resultado:** Efeito de uma intervenção ou organização sobre o seu público-alvo ou conjunto de destinatários diretos. São observáveis no momento da conclusão da intervenção (v.g. diminuição do tempo médio de deslocação resultante da conclusão de uma nova estrada). Este conceito pode assumir diferentes significados em diversas fontes de informação.

**Retorno econômico:** Segundo Menezes Filho et al. (2012) o retorno econômico compara o custo com o benefício financeiro gerado pelo programa. A avaliação de retorno econômico procura responder se o benefício gerado pelo projeto é maior do que seu custo; se o projeto é viável economicamente; e se existe algum projeto alternativo mais viável.

## S

**Seleção aleatória:** Dentro de um universo de indivíduos que possuam as mesmas características esperadas para participarem como público-alvo de um programa, porém os recursos disponíveis não são suficientes para atender a todos. A seleção aleatória é a divisão desse universo em dois grupos: tratamento e controle.

## T

**Taxa de Juros:** A taxa de juros quando mínima é a taxa mínima de atratividade que um investidor pretende ganhar com seu investimento.

**Taxa Interna de Retorno (TIR):** é a taxa que iguala custos e benefícios, ou seja, aquela que torna o valor presente líquido nulo. Essa taxa indica o retorno marginal do projeto, podendo ser comparada a de projetos alternativos ou com uma medida de custo de oportunidade para verificar se o projeto é viável.

**Top-down:** estratégias de processamento de informação e ordenação do conhecimento. Método pelo qual um investidor primeiramente observa as tendências na economia geral e em seguida escolhe os setores e as empresas que provavelmente se beneficiarão dessas tendências.

## V

**Validade externa:** assegura que os resultados na amostra sejam representativos dos resultados na população, dentro de um nível pré-definido de erro amostral.

**Validade interna:** o efeito observado sobre a variável dependente se deve ao tratamento e não a outros fatores.

**Valor Futuro:** É quanto determinado montante de dinheiro recebido no tempo inicial valerá em momentos posteriores.

**Valor Presente:** É quanto determinado montante de dinheiro a ser recebido no futuro vale no tempo inicial do investimento.

**Valor Presente Líquido:** É uma estatística de retorno econômico que mede a relação entre o custo e o benefício gerado pelo projeto. Portanto, o valor presente líquido estima o retorno econômico por meio da subtração do valor presente do custo total pelo valor presente do benefício total. Se os benefícios forem maiores que os custos, então o projeto é viável economicamente.

**Variáveis qualitativas:** indicam a presença ou a ausência de um atributo, ou “qualidade”.

**Variável dummy:** é um método para quantificar tais atributos a partir de variáveis artificiais que assumam valores de 1 ou 0 (0 indicando a ausência de um atributo e 1 a presença deste).

**Vazamento do programa:** incluir no programa pessoas não pertencentes ao público-alvo.

# Termos Estatísticos

## A

**Amostra:** Um subconjunto de observações selecionado numa população.

**Amostra por conveniência:** amostra não probabilística, é formada obedecendo a algum tipo de conveniência de quem forma a amostra ou de quem vai participar da amostra ou de ambos.

**Amplitude:** é a diferença entre o maior e o menor valor do conjunto de dados.

**Análise de Regressão:** análise da dependência de uma variável, a variável dependente, em relação a uma ou mais variáveis, as variáveis explanatórias, visando estimar e/ou prever o valor médio (da população) da primeira em termos dos valores conhecidos ou fixados (em amostragens repetidas) das segundas.

**Assimetria:** a distribuição apresenta valores diferentes para sua média, mediana e moda, dessa maneira, não coincidem no ponto central da distribuição.

## B

**Beta – convergência:** refere-se à convergência condicional das taxas de crescimento entre diferentes unidades econômicas (como países, regiões, estados, municípios ou setores)

## C

**Coeficiente de variação:** é uma medida de variabilidade relativa, definida como a razão percentual entre o desvio padrão e a média, e assim sendo uma medida adimensional expressa em percentual. Ele é usado para comparar a dispersão de diferentes conjuntos de dados, independentemente das unidades de medida das variáveis.

$$cv = \left( \frac{s}{\bar{x}} \right) \times 100$$

**Convergência:** É um conceito que sugere que economias menos desenvolvidas tendem a crescer mais rapidamente do que economias mais avançadas, permitindo que elas se aproximem do nível de renda dos países mais desenvolvidos ao longo do tempo.

**Convergência absoluta:** Refere-se a um fenômeno econômico no qual economias com níveis de renda mais baixos tendem a crescer mais rapidamente do que economias com níveis de renda mais altos, resultando em uma redução da disparidade de renda ao longo do tempo.

**Convergência condicional:** Significa que as economias mais pobres, que começam com níveis de renda mais baixos, têm uma tendência a crescer mais rapidamente do que as economias mais ricas, reduzindo assim as diferenças de renda ao longo do tempo.

**Coorte:** Estudo observacional que avaliou um grupo de pessoas com uma característica ou um conjunto de características em comum. Tipicamente o grupo é acompanhado durante um período específico para determinar a incidência de um distúrbio ou complicações (prognóstico).

**Correlação:** mede o grau pelo qual duas variáveis tendem a mudar juntas. O coeficiente descreve a força e a direção da relação. Um coeficiente de correlação é a covariância dividida pelo produto do desvio padrão de cada variável. O valor de  $\rho_{x,y}$  oscila no intervalo entre 0 e 1.

$$\text{corr}(x, y) = \rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}$$

**Covariância:** mede a relação linear entre duas variáveis. Os valores de covariância podem variar de infinito negativo a infinito positivo. A covariância será positiva se as duas variáveis tendem a variar no mesmo sentido, isto é, valores de X acima da sua média estão associados a valores de Y acima de sua média, o mesmo ocorrendo para valores de ambos, inferiores à média. A covariância será negativa se valores acima da média de uma variável estão associados a valores inferiores à média da outra.

$$\text{cov}(x, y) = E[(x - E(x)) - (y - E(y))]$$

**Critério de exclusão:** Critério que torna inelegíveis potenciais participantes de um estudo específico.

**Curtose:** Um indicador da espessura das caudas de uma distribuição baseada no quarto momento da variável aleatória padronizada; a medida é normalmente comparada com o valor da distribuição normal padrão, que é três.

## D

**Dados transversais (cross-section):** são dados temporais calculados para unidades mais agregadas tais como escolas, hospitais, setores de atividade, bairros, municípios ou estados.

**Decil:** É qualquer um dos nove valores que dividem os dados ordenados de uma variável em dez partes iguais, onde cada parte representa 1/10 da amostra ou população.

**Desvio médio:** É a diferença entre o valor observado e a medida de tendência central do conjunto de dados.

**Desvio padrão:** É a raiz quadrada da variância e sua unidade de medida é a mesma que a do conjunto de dados.

**Distribuição de Frequência:** Conjunto de dados divididos em classes pré-estabelecidas por certa frequência que serve de base para representações gráficas.

**Distribuição de Probabilidade:** É um modelo matemático que relaciona um certo valor da variável em estudo com a sua probabilidade de ocorrência. Há dois tipos de distribuição de probabilidade: Contínua e Discreta.

**Diferenças em diferenças:** Método baseado no cálculo de uma dupla subtração, a primeira se refere à diferença das médias da variável de resultado entre os períodos anterior e posterior ao programa, para o grupo de tratamento ( $\Delta Y_{T_1, t_1-t_0}$ ) e para o de controle ( $\Delta Y_{T_0, t_1-t_0}$ ), e a segunda se refere a diferença da primeira diferença calculada entre esses dois grupos. Denotando por  $T = \{1,0\}$  a participação ou não no programa e por  $t = \{1,0\}$  os períodos posterior e anterior a intervenção, respectivamente, o estimador de DD será dado por:  $\beta_{DD} = [E[Y|T = 1, t = 1] - E[Y|T = 1, t = 0}] - [E[Y|T = 0, t = 1] - E[Y|T = 0, t = 0]]$ .

**Diferenças em diferenças escalonado:** É uma abordagem estatística para análise de impacto causal que estende o método de diferenças em diferenças (DID) tradicional, onde os tratamentos podem ser implementados em diferentes momentos no tempo em várias unidades.

**Distribuição Normal:** Uma distribuição de probabilidades comumente usada em estatística e econometria para modelar uma população. Sua função de distribuição de probabilidades tem a forma de um sino.

**Distribuição T de Student:** A distribuição da razão de uma variável aleatória normal padrão e da raiz quadrada de uma variável aleatória independente qui-quadrada, na qual a variável aleatória qui-quadrada é primeiro dividida por seus  $gl$ .

## E

**Efeito calendário:** É o efeito médio de participar do tratamento em um período de tempo  $t$  (entre grupos que adotaram o tratamento até o período  $t$ ).

**Erro Padrão:** Avalia a precisão da média populacional. É a razão entre o desvio padrão e a raiz quadrada do tamanho da amostra.

**Erro Quadrático Médio:** A distância quadrada esperada que um estimador seja do valor populacional; ele iguala a variância mais o quadrado de qualquer viés.

**Erro Tipo I:** Rejeita a hipótese nula ( $H_0$ ) quando a hipótese nula é verdadeira.

**Erro Tipo II:** Não rejeita a hipótese nula ( $H_0$ ) quando a hipótese nula é falsa.

**Escore de Propensão:** a probabilidade estimada em relação a um conjunto de características que pode ser usada para definir grupos de estudos e avaliá-los, como por exemplo, selecionar grupos de tratamento e controle.

**Estatística F:** A estatística F é uma medida estatística utilizada em análises de variância (ANOVA) e em testes de significância em modelos de regressão. O teste F pode ser visualizado a partir da seguinte formula:  $F = \frac{SQE/gl}{SQR/gl} = \frac{SQE/(k-1)}{SQR/(n-k)}$ . Assim, se  $F > F_\alpha(k-1, n-k)$ , rejeita-se  $H_0$ , caso contrário, não rejeite, em que  $F_\alpha(k-1, n-k)$  é o valor crítico de F no nível  $\alpha$  de significância;  $(k - 1)$ , os graus de liberdade do numerador, e  $(n - k)$ , os graus de liberdade do denominador.

**Estatística t:** É um teste de hipótese que segue uma distribuição t de student. Utiliza-se de conceitos estatísticos para rejeitar ou não uma hipótese nula. Fórmula do teste:  $t = \frac{\hat{\beta}_2 - \beta}{ep(\hat{\beta}_2)}$ , onde  $\hat{\beta}_2$  é o valor estimado do parâmetro e  $\beta$  o valor do parâmetro sob a hipótese nula.

**Estimador:** Uma regra de combinar dados para produzir um valor numérico de um parâmetro populacional; a forma da regra não depende da amostra particular obtida.

**Estudo de eventos:** O modelo de estudo de eventos é uma ferramenta econométrica usada para estimar os efeitos do tratamento dinâmico.

**Função Densidade de Probabilidade:** É uma função que descreve a probabilidade relativa de uma variável aleatória tomar um valor dado.

## H

**Heterogeneidade:** individualidade ou originalidade de diferentes indivíduos de uma amostra.

**Hipótese:** é uma explicação proposta para um fenômeno observável.

**Hipótese Alternativa:** Em avaliação de impacto, a hipótese alternativa é geralmente a suposição de que a hipótese nula é falsa; em outras palavras, que a intervenção possui um impacto nos resultados.

**Hipótese Nula:** A hipótese nula é uma hipótese que pode ser falsificada com base nos dados observados. A hipótese nula tipicamente propõe uma posição geral ou padrão. Em avaliação de impacto, a posição padrão geralmente é que não há diferença entre o grupo de tratamento e o grupo de controle, ou em outras palavras, que a intervenção não possui nenhum impacto nos resultados.

**Homogeneidade:** quando os dados associados às categorias de uma das variáveis se comportam de modo homogêneo ou similar nas diversas classes ou populações definidas pelas categorias da outra variável classificatória.

**Heterocedasticidade:** A variância do termo de erro, dadas as variáveis explicativas, não é constante.

**Homocedasticidade:** Os erros em um modelo de regressão têm variância constante, condicional nas variáveis explicativas.

**Inferência estatística:** refere-se ao uso apropriado dos dados da amostra para se ter algum conhecimento sobre os parâmetros da população.

**Intervalo de Confiança:** procura captar os possíveis valores que o parâmetro poderia assumir se amostras diferentes fossem utilizadas, com uma probabilidade de certeza.

## L

**Lei do Grandes Números (LGN):** Um teorema que diz que a média de uma amostra aleatória converge em probabilidade para a média populacional, a LGN também é válida para séries temporais estacionária e fracamente dependentes.

## M

**Média aritmética:** é a soma de todas as observações dividida pelo número de observações.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

**Mediana:** é uma medida de posição definida como o número que se encontra no centro de uma série de números, estando estes dispostos segundo uma ordem crescente ou decrescente.

**Moda:** é o valor que ocorre com maior frequência em uma série de valores

**Mínimos Quadrados Ordinários (MQO):** O método dos mínimos quadrados é uma abordagem estatística usada para determinar a melhor aproximação linear entre um conjunto de dados e um modelo matemático. Seu propósito fundamental é identificar os parâmetros do modelo que reduzem ao mínimo a soma dos quadrados das discrepâncias entre os valores observados e os valores previstos pelo modelo.

## N

**Nível de significância:** A probabilidade de erro tipo I em testes de hipóteses.

## O

**Outlier:** ocorre quando um dado se diferencia drasticamente dos demais. Em outras palavras, um outlier é um valor que foge da

normalidade e que pode (e, provavelmente, deve) causar anomalias nos resultados obtidos por meio de algoritmos e sistemas de análise.

## P

**Padrão ouro:** É o método diagnóstico padrão disponível no momento que dará maior grau de certeza para classificar corretamente a amostra.

**Parâmetro:** Um valor desconhecido que descreve uma relação populacional.

**Percentil:** Percentil são os noventa e nove valores que separam uma série em 100 partes iguais.

**Poder estatístico:** Em uma comparação de duas intervenções, o poder é a habilidade de detectar uma diferença entre as duas, se essa diferença realmente existir.

**População:** O grupo de unidades que são elegíveis para receber uma intervenção ou tratamento.

**P-valor:** é a probabilidade de que a estatística do teste tenha valor extremo em relação ao valor observado (estatística) quando a hipótese nula é verdadeira. Portanto, é o nível de significância observado ou exato, ou seja, é o nível mínimo de significância ao qual a hipótese nula pode ser rejeitada (probabilidade de cometer o erro tipo I).

**Poder do teste:** A probabilidade de rejeitar a hipótese nula quando ela for falsa; o poder depende dos valores dos parâmetros populacionais sob a hipótese alternativa.

## Q

**Quartil:** os valores de uma série que a dividem em quatro partes iguais.

**Quintil:** os valores de uma série que a dividem em cinco partes iguais.

## R

**R<sup>2</sup>:** Medida da qualidade do ajustamento de um modelo de Análise de Regressão. Seus valores estão entre 0 e 1, quanto mais próximo de 1 melhor o ajustamento. A fórmula para o R<sup>2</sup> é:

$$R^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}$$

**R<sup>2</sup> ajustado:** O coeficiente de determinação ajustado (R<sup>2</sup> ajustado) é uma métrica estatística utilizada na análise de regressão para avaliar a qualidade de um modelo de regressão, levando em consideração o número de variáveis independentes incluídas no modelo. Diferente do R<sup>2</sup>, o R<sup>2</sup> ajustado pode ser negativo, Enquanto o R<sup>2</sup> (coeficiente de determinação) mede a proporção da variância na variável dependente que é explicada pelas variáveis independentes no modelo, o R<sup>2</sup> ajustado corrige o R<sup>2</sup> para penalizar a inclusão de variáveis independentes desnecessárias ou redundantes no modelo. A fórmula para o R<sup>2</sup> ajustado é:

$$\bar{R}^2 = 1 - \frac{SQR/n - k}{SQT/n - 1} = 1 - (1 - R^2) \frac{n - 1}{n - k}$$

**Regressão simples:** análise da dependência de uma variável em relação a uma única variável explanatória.

**Regressão múltipla:** análise da dependência de uma variável em relação a mais de uma variável explanatória.

## S

**Seleção aleatória:** método de aleatorização que permite criar uma situação na qual não há correlação entre ser ou não tratado e os atributos das unidades de observação. Fornece o balanceamento necessário tanto das características observadas quanto das não observadas das unidades que compõem os dois grupos, eliminando dessa forma, o viés de autosseleção.

**Série temporal:** conjunto de observações dos valores que uma variável assume em diferentes momentos do tempo.

**Sigma-convergência:** este tipo de convergência acontece quando é observada uma queda na dispersão do produto per capita de um conjunto de países ou regiões dentro de um dado intervalo de tempo.

**Significância Estatística:** está relacionada ao nível de confiança ao rejeitar a hipótese nula quando ela é verdadeira (erro do tipo I).

## T

**Tamanho amostral:** número de observações contidas na amostra.

**Teste de Hipótese:** Método de inferência estatística que analisa uma amostra a partir da probabilidade. São constituídos de alternativas que serão testadas.

**Teste do Qui-quadrado:** Teste estatístico que examina a distribuição dos desfechos categóricos em dois grupos. A hipótese nula é de que as distribuições subjacentes são idênticas.

**Teorema do Limite Central:** Um resultado básico da teoria da probabilidade implica que a soma das variáveis aleatórias independentes, ou mesmo das variáveis aleatórias fracamente dependentes, quando padronizadas por seu desvio-padrão tem uma distribuição que tende a normal-padrão quando o tamanho da amostra cresce.

**Termo de erro aleatório:** a variável em uma equação de regressão simples ou múltipla que contém fatores não observados que afetam a variável dependente. O termo de erro também pode incluir erros de medida nas variáveis dependentes ou independentes observadas.

## U

**Universo:** conjunto de elementos sobre o qual incide o estudo estatístico.

## V

**Valor absoluto:** valor do próprio algarismo, independentemente do lugar que ocupa no número.

**Variância:** é uma medida que expressa um desvio quadrático médio do conjunto de dados, e sua unidade é o quadrado da unidade dos dados. Em outras palavras, a variância indica o quanto distantes os valores de uma variável estão da média.

$$var = s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

**Variável aleatória:** É uma variável quantitativa. Uma função  $X$  que associa a cada elemento  $W$  do espaço amostral um valor  $X$  pertencente ao conjunto dos reais.

**Variável aleatória Contínua:** Uma variável aleatória contínua é um tipo de variável aleatória  $X$  que pode assumir um número infinito de valores dentro de um intervalo contínuo de números reais.

**Variável aleatória Discreta:** Dado que X é uma variável aleatória. Se o número de valores possíveis de X for enumerável (finito ou infinito), dizemos que X é uma variável aleatória discreta. Portanto, os possíveis valores de podem ser ordenados como,  $X^1 \ X^2$ .

**Variável dummy:** variável que assume valores zero ou um.

**Variável exógena:** qualquer variável que seja não correlacionada com o termo de erro no modelo de interesse.

**Variável endógena:** variável que é correlacionada com o termo de erro no modelo de interesse.

**Variável instrumental:** em uma equação com uma variável explicativa endógena, uma variável instrumental é uma variável que não aparece na equação, é não correlacionada com o erro na equação, e é (parcialmente) correlacionada com a variável explicativa endógena.

**Variável proxy:** uma variável observada que é relacionada, mas não idêntica a uma variável explicativa não observada em análise de regressão múltipla.

**Viés:** a diferença entre o valor esperado de um estimador e o valor populacional que o estimador supostamente está estimando.

**Viés de seleção amostral:** viés do estimador de Mínimos Quadrados Ordinários (MQO) que é induzido pelo uso de dados surgidos da seleção amostral endógena.

# Indicadores Econômicos e Sociais

## C

**Crescimento econômico:** Aumento da capacidade produtiva da economia e, portanto, da produção de bens e serviços de determinado país ou área econômica.

**Capital Físico:** pode ser definido como a quantidade de ativos não-humanos empregados na produção.

**Capital Humano:** conjunto de habilidades e conhecimentos capazes de tornar os trabalhadores mais produtivos e se acumulam através do investimento em educação.

## D

**Desigualdade de renda:** Má distribuição de renda que ocasiona um desequilíbrio econômico entre uma população, estando a maior concentração de riquezas nas mãos de uma minoria; desigualdade econômica.

## E

**Empregabilidade:** conjunto de competências e habilidades necessárias para uma pessoa conquistar e manter um trabalho.

## G

**Grau de Abertura (GA):** razão entre a soma de exportações e importações e o PIB.

## H

**Hiato da Pobreza:** fornece uma medida da intensidade média da pobreza para as famílias que estão abaixo da linha de pobreza.

$$P = \frac{1}{qz^*} \sum_{i=1}^Q (z^* - y_i)$$

em que  $P$  é o hiato de pobreza,  $z^*$  é a linha de pobreza,  $Q$  é o conjunto de todas as famílias pobres. O hiato (gap) entre a linha de pobreza e a renda per capita de uma família  $i$  abaixo da linha de pobreza ( $z^* - y_i$ ) mede a insuficiência de renda dessa família em relação à linha de pobreza.

**Índice de Gini:** É uma medida de concentração ou desigualdade geralmente utilizada na análise da distribuição de renda, podendo ser utilizada também para medir o grau de concentração de qualquer distribuição estatística. Por exemplo: grau de distribuição da população urbana de um país, por município. Varia de zero (igualdade perfeita) a 1 (desigualdade máxima).

**Índice de Pobreza Humana (IPH):** índice criado em 1997 pelo PNUD para medir o grau de pobreza humana ou de privação de meios básicos à sobrevivência. Enquanto o IDH mede os progressos médios nas dimensões de educação, saúde e recursos materiais para sobrevivência, o IPH avaliará o nível de privação nestas dimensões.

**Índice de preço:** Instrumento de avaliação de preços que mede, em um dado período, em uma determinada população, as variações percentuais registradas nos preços de um conjunto de bens e serviços de diferentes setores, tais como: alimentação, bebidas, vestuário, moradia, saúde, transporte, dentre outros.

**Índice de Theil:** O Índice de Theil é uma medida de desigualdade que avalia a distribuição de uma variável em uma população. Ele mede a dispersão da distribuição de renda ou riqueza entre os indivíduos. Em essência, o Índice de Theil varia de 0 a 1, onde 0 representa igualdade perfeita (ou ausência de desigualdade) e 1 representa desigualdade máxima (ou maior desigualdade possível) na distribuição da variável analisada. Quanto mais próximo o índice estiver de 1, maior é a desigualdade na distribuição da variável em questão.

## M

**Massa salarial:** A massa salarial é o total de remuneração paga a todos os trabalhadores de uma empresa, setor econômico ou país em um período específico, como um mês ou ano. Isso inclui os salários e benefícios recebidos pelos trabalhadores por seu trabalho, abrangendo toda a força de trabalho ativa considerada na análise. Assim, a massa salarial é essencial para compreender a dinâmica econômica e social, pois reflete o volume de renda gerado pelos trabalhadores em uma economia ou organização.

# P

**Peso do Governo:** razão entre a receita tributária e o Produto Interno Bruto.

**Pobreza:** Estado de carência em que vivem indivíduos ou grupos populacionais, impossibilitados, por insuficiência de rendas ou inexistência de bens de consumo, de satisfazer suas necessidades básicas de alimentação, moradia, vestuário, saúde e educação.

**Pobreza extrema:** de acordo com o Banco Mundial, é uma situação em que uma pessoa vive com uma renda menor que US\$ 1,90 por dia. Também é considerado, população com renda média domiciliar per capita menor que  $\frac{1}{4}$  de salário-mínimo.

**Proporção de pobres:** Percentual da população residente com renda familiar mensal per capita de até meio salário-mínimo, em determinado espaço geográfico, no ano considerado.

# R

**Renda Média:** Média das rendas domiciliares per capita das pessoas residentes em determinado espaço geográfico, no ano considerado.

# S

**Salário:** Remuneração recebida por serviço(s) prestado(s); remuneração que corresponde ao que foi estabelecido como pagamento no contrato de trabalho; ordenado.

# V

**VAB Serviços:** É uma medida econômica que representa o valor agregado por todas as atividades de serviços em uma economia. O VAB de Serviços é calculado subtraindo-se o valor dos insumos utilizados na produção de serviços do valor total gerado por essas atividades.

**VAB Indústria:** É uma medida econômica que representa o valor agregado por todas as atividades industriais em uma economia. O VAB Indústria é calculado subtraindo-se o valor dos insumos utilizados na produção de bens do valor total gerado por essas atividades.

**VAB Agricultura:** É uma medida econômica que representa o valor agregado por todas as atividades agrícolas em uma economia. O VAB

Agricultura é calculado subtraindo-se o valor dos insumos utilizados na produção agropecuária do valor total gerado por essas atividades.

**Vínculos:** Referem-se aos contratos ou relações formais entre empregadores e funcionários.

# Fontes Externas de Dados

BEN: Balanço Energético Nacional.

Variáveis construídas a partir das informações do BEN:

Capital Físico Estadual

Capital Físico Municipal

IBGE: Instituto Brasileiro de Geografia e Estatística.

Variáveis construídas a partir das informações do IBGE:

Crescimento Econômico

Desigualdade de renda

Empregabilidade

Hiato da Pobreza

Índice de Gini

Índice de Preço

Índice de Theil

Massa Salarial Estadual

Pobreza

Pobreza Extrema

Proporção de pobres

Renda Média Estadual

Salário

VAB Serviços

VAB Indústria

VAB Agricultura.

MDIC: Ministério da Indústria, Comércio Exterior e Serviços.

Variáveis construídas a partir de informações do MDIC:

Grau de Abertura.

MTE: Ministério do Trabalho e Emprego.

RAIS: Relação Anual de Informações Sociais.

Variáveis construídas a partir das informações do RAIS:

Vínculos

Renda Média Municipal

Capital Humano Municipal

Capital Humano Estadual.

SIBF: Sistema de Informações dos Benefícios e Incentivos Fiscais.

Variáveis construídas a partir das informações do SIBF:

Número de pleitos de incentivos aprovados.

STN: Secretaria do Tesouro Nacional.

Variáveis construídas a partir das informações da STN:

Peso do governo.