

Econometria com Dados em Paineis: Fundamentos e Aplicações

Adriano Vargas Saldanha

Mestrando no Programa de Pós-Graduação em
Organizações e Mercados - UFPel

Pelotas/RS, novembro de 2025



Índice

- Apresentação
- Objetivos
- O que é Econometria?
- Modelo Econométrico vs Econômico
- Cross-Section
- Séries Temporais
- Dados em Painel
- Instalação do VSCode e integração com Python

Apresentação



→ Formação

- Técnico em Edificações – IFSul/Pelotas
- Graduação em Ciências Econômicas – UFPel
- Mestrando em Economia Aplicada - PPGOM/UFPel

→ Linhas de pesquisa

- Economia da Saúde
- Economia do Bem-Estar

Objetivos

→ Objetivo Geral

- Capacitar o participante a compreender e aplicar os fundamentos da econometria com dados em painel, desde os conceitos básicos de modelagem econômica até a estimação e interpretação de modelos, com aplicação prática no software Python.

Objetivos

→ Objetivos Específicos

- Introduzir conceitos fundamentais de modelagem econômica e sua relevância para análise empírica;
- Apresentar os tipos de dados utilizados em economia: Corte Transversal (Cross-Section), Séries Temporais e Painel;
- Ensinar a organizar e estruturar bases de dados para análise;
- Estimar modelos básicos de Dados em Painel (Pooled OLS, Efeitos Fixos e Efeitos Aleatórios); e
- Desenvolver habilidades para interpretar e comunicar os resultados

O que é Econometria?

→ Definição

- Segundo Woodridge (2013), Econometria é o conjunto de métodos estatísticos usados para estimar, testar e avaliar relações econômicas teóricas por meio de dados observados;
- No geral, a Econometria mistura em suas análises três conhecimentos distintos, sendo a Estatística, a Matemática e a Teoria Econômica;
- Atualmente, a Econometria também passou a agregar outras fontes auxiliares de análise, como o uso de softwares para rodar modelos, realizar inferências, dentre outras coisas.

O que é Econometria?

→ Como funciona?

- A Econometria em seu campo de atuação possui várias ramificações de trabalho, sendo utilizada também para avaliação de políticas econômicas, Data Science, Planejamento Econômico, dentre outras áreas do conhecimento;
- Seu funcionamento é voltado para analisar uma série de dados e conectar um modelo teórico com todos os dados obtidos a fim de quantificar a teoria e checar como aquela pressuposição teórica pode ser aplicada na realidade;
- É praticamente um meio de verificabilidade das teorias econômicas existentes, quantificando-as na tentativa de solucionar problemas econômicos.

O que é Econometria?

→ Exemplo Prático:

- A Econometria pode utilizar modelos econométricos (os quais quantificam uma relação teórica) para responder as seguintes questões:
 - O quanto um aumento de 15% no valor do produto vendido pode impactar no valor das vendas?
 - O quanto uma elevação no nível de renda pessoal pode gerar no aumento do valor poupado?
 - O quanto o aumento no preço do X-Salada pode influenciar no preço do X-Bacon?

Modelo Econométrico vs Modelo Econômico

→ Modelo Econométrico

- Quantifica a relação entre as variáveis em um modelo, possibilitando que sejam feitas previsões e análises de quanto uma variável impacta em outra variável;
- Utiliza ferramental estatístico, matemático e outros meios para explicar uma relação teórica e medir sua aplicabilidade;
- Necessita de dados para ser executado, já que sem os dados é impossível fazer uma análise de previsão. Logo, não analisa a relação entre as variáveis, mas mensura o quanto, em proporção, uma variável afeta a outra.

Modelo Econométrico vs Modelo Econômico

→ Modelo Econométrico

- Componentes de um modelo econométrico:
 - Variável dependente (Y): o que queremos explicar;
 - Variável independente (X): fatores que influenciam Y ;
 - Parâmetros (β): medem a intensidade e direção do efeito;
 - Termo de erro (ε): fatores não observados
- Um modelo base pode ser descrito como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon$$

Modelo Econométrico vs Modelo Econômico

→ Modelo Econômico

- Busca trazer uma relação entre as variáveis. Estuda o comportamento e, de maneira teórica, mostra como uma variável se relaciona com a outra;
- O Modelo Econômico não é quantificado, de modo que fica restrito a sua parte teórica. Apenas indica o que pode ocorrer na realidade, mas não verifica através da mensuração de dados;
- Não descarta o uso de conhecimentos adicionais. No entanto é incapaz de fazer previsões através de quantificação, já que precisa ser testado para haver verificabilidade.

Modelo Econométrico vs Modelo Econômico

→ Ex. Modelo Econômico

- Aqui possuímos um modelo que relaciona o preço com a quantidade demandada. Entretanto, essa relação precisa ser verificada por meio de dados e modelos econométricos, na tentativa de resolver o problema econômico:

$$Q_d = f(P, P_c, P_s, E, R, \dots)$$

- Essa função mostra que a quantidade demandada é função do preço, preço de bens complementares, preço de bens substitutos e outros fatores;
- No entanto a mesma função não tem a capacidade de fornecer resultados factuais, como por exemplo números. Logo, precisamos de um modelo econométrico para analisar a situação, com base nesse modelo econômico.

Modelo Econométrico vs Modelo Econômico

→ Ex. Modelo Econométrico

- O modelo econométrico é o principal mecanismo utilizado pela Econometria para realizar previsões. Contudo, a criação de um modelo econométrico depende do modelo teórico, de modo que possamos utilizar o ferramental estatístico e matemático para resolver problemas práticos;
- Logo, vamos supor um modelo econométrico que vai utilizar a base da função demanda anterior para resolver o seguinte caso:

Modelo Econométrico vs Modelo Econômico

→ Ex. Modelo Econométrico

- Um dono de uma cervejaria quer saber o quanto o aumento de R\$1,00 na lata de cerveja vai impactar na sua quantidade demandada de cervejas, e quanto um acréscimo de R\$1,00 na cerveja sem álcool afeta as vendas da cerveja com álcool. Sendo assim, tem-se o seguinte modelo econométrico, exemplificado por um modelo de regressão linear múltipla:

$$\text{Quantidade Vendida de Cerveja}_t = \beta_0 + \beta_1 (\text{Preço Cerveja}_t) + \beta_2 (\text{Preço Cerveja SA}_t) + \varepsilon_t$$

- Esse modelo sendo estimado e com sua análise posterior de inferência, tem a capacidade de responder dúvidas do dono do bar.

Dados em Corte Transversal

→ Definição

- Os Dados em Corte Transversal ou Cross-Section são um conjunto de dados retirados em um instante de tempo (não há variações no tempo), realizando várias análises, como por exemplo, um conjunto de alunos de uma classe ou um determinado grupo de jogadores de um time;
- Quando trabalhamos com Corte Transversal, temos diferentes unidades de análise que apresentam seus respectivos comportamentos específicos. Cada unidade se comportará de maneira diferente.

Dados em Corte Transversal

→ Características

- Uma série de corte transversal pode ser vista nesta tabela ao lado, onde temos dados de consumo mensal e renda mensal de 12 alunos em março de 2024. Nesse caso, temos 12 unidades de análise diferentes (12 alunos), cada uma com seu respectivo comportamento, e temos um corte no tempo sem variação (Março de 2024). Assim, temos uma série de Corte Transversal.

| Renda e Consumo dos Alunos da Turma em Março de 2024 | | |
|--|---------------|-------------|
| Aluno | Consumo (R\$) | Renda (R\$) |
| Adriano | 52,25 | 258,30 |
| Bruno | 58,32 | 343,10 |
| Carla | 81,79 | 425,00 |
| Elissa | 119,90 | 267,50 |
| Felipe | 100,46 | 487,70 |
| Gabriela | 121,51 | 496,50 |
| Isabel | 98,48 | 588,30 |
| João | 181,21 | 591,30 |
| Marcos | 107,48 | 564,60 |
| Patrícia | 92,84 | 631,00 |
| Thiago | 104,94 | 548,70 |
| Xuxa | 129,57 | 611,20 |

Dados em Corte Transversal

→ Modelo Econométrico

- Em uma Regressão Linear Simples com dados retirados de corte transversal, temos um i para representar o momento do corte:

$$\text{Despesa}_i = \beta_0 + \beta_1(\text{Renda}_i) + \varepsilon_i$$

- Percebam que o i mostra que esses dados do modelo foram tirados de uma série de corte transversal. Nesse sentido, o i é colocado na frente da variável dependente, explicativa e o erro.

Dados em Corte Transversal

→ Vantagens

- Simplicidade: mais fácil de coletar e trabalhar para análises iniciais;
- Fotografia do momento: útil para comparações entre unidades no mesmo período;
- Disponibilidade: muitos órgãos e pesquisas divulgam dados anuais e corte;
- Custo: geralmente mais barato e rápido de obter que séries temporais longas ou painéis.

Dados em Corte Transversal

→ Desvantagens

- Não captura dinâmica temporal: não mostra como variáveis evoluem ao longo do tempo;
- Causalidade limitada: maior risco de viés por omissão ou simultaneidade;
- Heterogeneidade não observada: características fixas não medidas podem distorcer resultados;
- Momentâneo: conclusões podem não valer para outros períodos.

Séries Temporais

→ Definição

- Segundo Wooldridge, uma série temporal é um conjunto de observações sobre uma variável (ou múltiplas variáveis) coletadas ao longo do tempo em intervalos regulares;
- Diferentemente dos dados de corte transversal, onde a ordem das observações é irrelevante, nas séries temporais a ordenação cronológica é essencial para a análise;
- A natureza temporal dos dados cria dependências entre observações consecutivas, exigindo técnicas econométricas específicas que reconheçam essa estrutura.

Séries Temporais

→ Características

- Dependência Temporal:
 - As observações de uma série são correlacionadas ao longo do tempo; o valor atual costuma depender de valores passados. Por isso, são usados modelos específicos como AR, MA, ARMA e ARIMA.
- Estacionaridade:
 - Uma série é estacionária quando suas propriedades básicas (média, variância e covariância) permanecem constantes ao longo do tempo.

Séries Temporais

→ Características

- Periodicidade:
 - As séries variam conforme a frequência da coleta, podendo ser diárias, mensais, trimestrais ou anuais. A periodicidade determina o tipo de dinâmica e modelo apropriado (ex.: inflação mensal vs. PIB anual).
- Choques e estrutura dinâmica:
 - Variáveis econômicas são afetadas por choques (mudanças inesperadas). A econometria de séries temporais busca entender como esses choques se propagam e influenciam o tempo.

Séries Temporais

→ Características

- Uma série temporal **Trabalha com período de tempo e não com instante de tempo. O tempo varia, ao modo que temos um período de tempo;**
- A tabela ao lado analisa os níveis de consumo e renda do Adriano entre janeiro de 2024 e dezembro de 2024. Essa é uma série temporal, pois tem uma única unidade de análise e trabalha num período de tempo.

| Níveis de Renda e Consumo do Adriano entre 01/24 e 12/24 | | |
|--|---------------|-------------|
| Período | Consumo (R\$) | Renda (R\$) |
| jan/2024 | 52,25 | 258,30 |
| fev/2024 | 58,32 | 343,10 |
| mar/2024 | 81,79 | 425,00 |
| abr/2024 | 119,90 | 267,50 |
| mai/2024 | 100,46 | 487,70 |
| jun/2024 | 121,51 | 496,50 |
| jul/2024 | 98,48 | 588,30 |
| ago/2024 | 181,21 | 591,30 |
| set/2024 | 107,48 | 564,60 |
| out/2024 | 92,84 | 631,00 |
| nov/2024 | 104,94 | 548,70 |
| dez/2024 | 129,57 | 611,20 |

Séries Temporais

→ Modelo Econométrico

- Quando utilizamos um modelo de regressão linear simples, devemos deixar explícito de onde ou qual tipo de série tiramos nossos dados. Assim, cada série de dados possui uma representação específica, onde a série temporal por sua vez é representada por um t .

$$\text{Despesa}_t = \beta_0 + \beta_1(\text{Renda}_t) + \varepsilon_t$$

- Reparem que esse t deve ficar na frente da variável dependente, explicativa e do erro, demonstrando que os dados desse modelo são provenientes de uma série temporal, onde estão analisando a despesa em função da renda de uma pessoa em um determinado período.

Séries Temporais

→ Vantagens

- Captura dinâmicas e tendências ao longo do tempo;
- Útil para previsão e avaliação de impactos de políticas;
- Pode incorporar defasagens e sazonalidade.

Séries Temporais

→ Desvantagens

- Precisa de histórico longo e dados de boa qualidade;
- Mais sensível a quebras estruturais;
- Modelagem e inferência mais complexas (estacionaridade, autocorrelação)

Dados em Painel

→ Definição

- **Dados em Painel (ou Dados Longitudinais)** são aqueles que contêm observações de múltiplas unidades (indivíduos, empresas, regiões, países etc.), acompanhadas ao longo de mais de um período de tempo;
- Em outras palavras, os Dados em Painel configuram uma série de dados mais complexa do que as demais séries, tendo um processo de rodagem de regressão diferenciado;
- Basicamente podemos descrevê-lo como a fusão entre uma série temporal com um corte transversal.

Dados em Painel

→ Características

- Combina informações de diferentes unidades (indivíduos, empresas, países) ao longo do tempo;
- Permite controlar a heterogeneidade não observada, reduzindo o viés de omissão de variáveis;
- Aumenta o número de observações, melhorando a eficiência dos estimadores;

Dados em Painel

→ Características

- Possibilita estudar as dinâmicas de ajuste, como impacto de políticas públicas econômicas no curto e longo prazo;
- Aplicações comuns:
 - Análise de desempenho de empresas ao longo dos anos;
 - Efeitos de políticas públicas em diferentes municípios;
 - Estudo de relação de risco-retorno em carteiras de ativos.

Dados em Painel

→ Características

- A tabela ao lado analisa a renda de três alunos de um curso de Economia. Aqui listamos os seus níveis de Renda e Poupança de 2021 até 2024;
- Logo, como possuímos um conjunto de unidades de análise em um período de tempo podemos dizer temos dados em painel. Ou seja, há variação de tempo e três unidades para análise.

| Análise dos Alunos de Economia - 2022/2023 | | | |
|--|------|----------|-------|
| Aluno | Ano | Poupança | Renda |
| Adriano | 2021 | 201 | 1452 |
| Adriano | 2022 | 321 | 1562 |
| Adriano | 2023 | 420 | 1600 |
| Adriano | 2024 | 423 | 1642 |
| Bruno | 2021 | 495 | 1468 |
| Bruno | 2022 | 524 | 1552 |
| Bruno | 2023 | 600 | 1600 |
| Bruno | 2024 | 574 | 1632 |
| Carla | 2021 | 503 | 1495 |
| Carla | 2022 | 589 | 1650 |
| Carla | 2023 | 623 | 1720 |
| Carla | 2024 | 620 | 1888 |

Dados em Painel

→ Modelo Econométrico

- Em modelos de regressão, os dados em painel são representados por it , simbolizando a mistura de série temporal e corte transversal.

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it}$$

- Desse modo, podemos considerar que os modelos de regressão com dados em painel são a união entre o i do corte transversal com o t da série temporal. Logo, sempre que se trabalha com dados em painel, é necessário fazer o uso do subscrito it .

Dados em Painel vs Corte Transversal

→ Corte Transversal

- Observações de várias unidades em um único ponto no tempo.

→ Dados em Painel

- Observações de várias unidades ao longo de múltiplos períodos de tempo.

→ Exemplo:

- Corte Transversal: Pesquisa de renda das famílias em 2024;
- Dados em Painel: Renda das mesmas famílias acompanhadas de 2020 a 2024.

Dados em Painel vs Séries Temporais

→ Séries Temporais

- Observações de uma única unidade ao longo do tempo.

→ Dados em Painel

- Observações de múltiplas unidades ao longo do tempo.

→ Exemplo:

- Séries de Tempo: Taxa Selic mensal de 2000 a 2025;
- Dados em Painel: Taxa de juros de vários países de 2000 a 2025.

Instalação do VSCode e integração com Python

→Link

<https://www.youtube.com/watch?v=R9dLGLVqK9Q&t=156s>