**DRAFT SOBRE MODELO DE REGRESSÃO EM DADO EM PAINEL**

**CLIENTE: Carlos Benassuly Maués Filho**

**AUTHOR: Mário Diego Rocha Valente**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Será apresentado uma revisão bibliográfica acerca do Modelo Estatístico Multivariado chamado Regressão de Dado em Painel. Descrevendo a formulação matemática, abordando os principais modelos de Dados Empilhados (Pooled OLS), Modelos de Efeitos Fixos (FE) e Modelos de Efeitos Aleatórios (RE). Ao final é apresentado o script criado na linguagem de Programação R.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Inicialmente, aplicou-se a análise fatorial a partir da matriz de correlações, no caso a base de dados é composta por 12 variáveis originais.

**Tabela 1**. Variáveis Originais utilizadas para Realização da Análise Fatorial Exploratória.

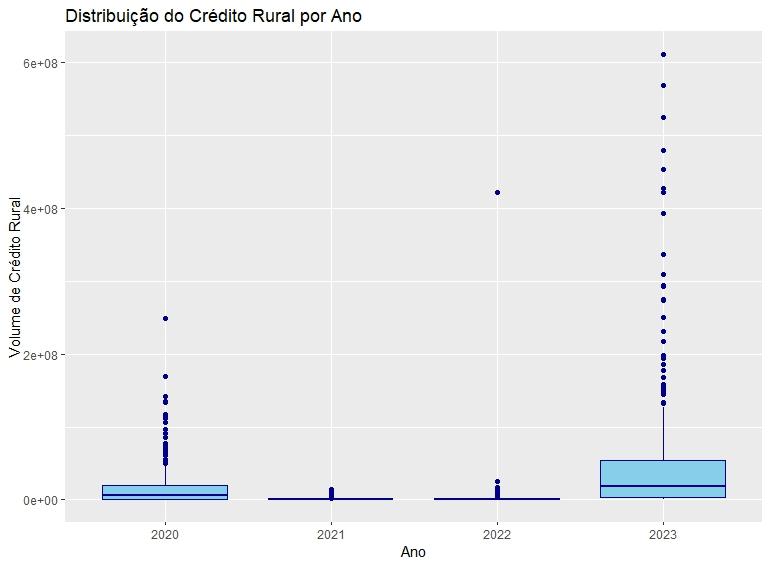
|  |  |
| --- | --- |
| **Itens** | **Variáveis** |
| Var1 | População |
| Var2 | Nº de Empregos Formais |
| Var3 | PIB(por mil) |
| Var4 | PIB Per Capita (R$ por pessoa) |
| Var5 | Nº de Agências Bancárias |
| Var6 | Inverso do Incremento do Desmatamento |
| Var7 | Nº de Estabelecimentos Agrícolas |
| Var8 | Rebanho Bovino (Cabeças) |
| Var 9 | Área Colhida (LP)(hectare) |
| Var10 | Área Colhida (LT)(hectare) |
| Var11 | Educação Básica |
| Var 12 | ICMS |

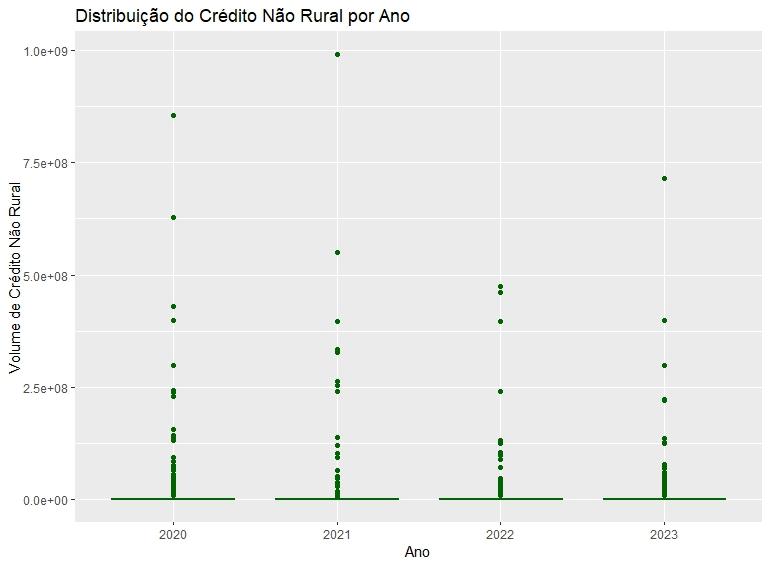
Aplicou-se a Análise Fatorial utilizando o método das componentes principais para extração dos fatores via rotação ortogonal do tipo varimax para obter melhores combinações e usando como critérios para a escolha do número de fatores a extrair, o Critério *Scree Test*, Critério da Raiz Latente e Critério da Percentagem da Variância Explicada. Assim, foi gerado um modelo com 3 fatores que explica 80,85% da estrutura de covariância inicial dos dados referentes ao Volume de Crédito Rural.

Sendo que o primeiro fator explica **52,22%**, o segundo fator explica **17,37%**, e o terceiro fator explica **11,26%**. Com isso, buscou-se identificar as variáveis que mais influenciam em cada fator, ou seja, as que possuem maiores cargas fatoriais.

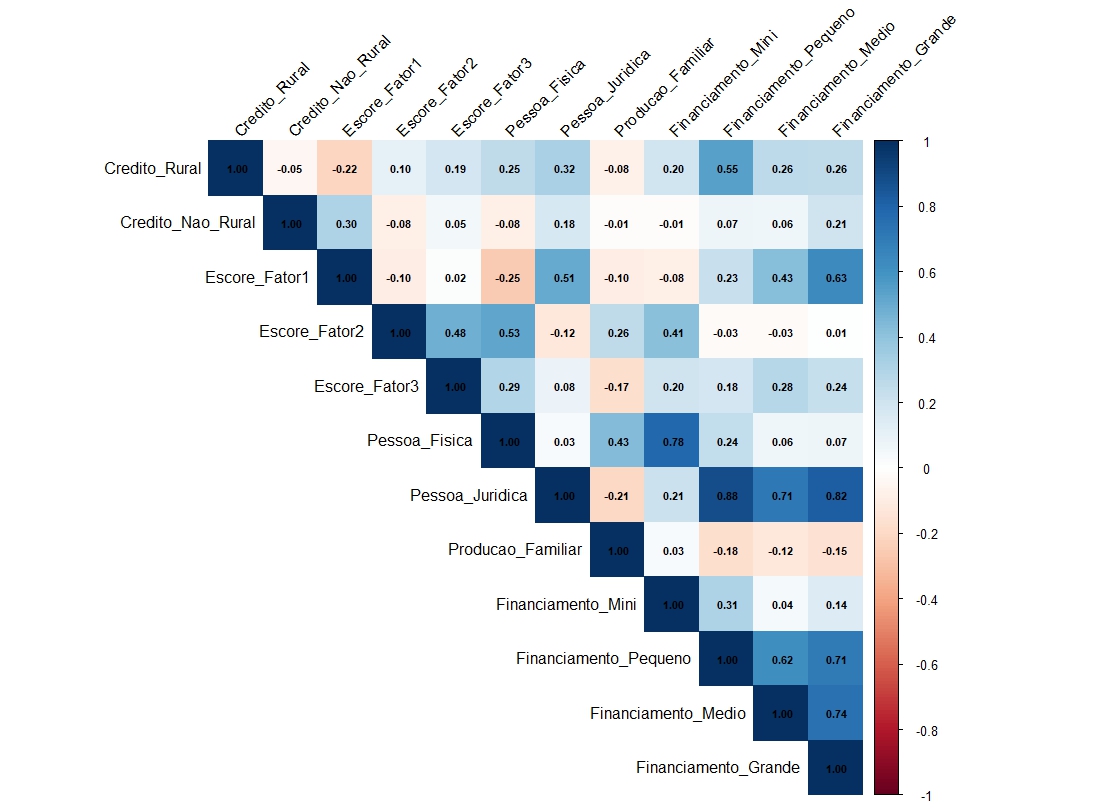
Com os resultados obtidos, pode-se observar que o primeiro fator possui pesos mais altos nas variáveis: População, Nº de Empregos Formais, Nº de Agências Bancárias e Educação Básica. O segundo fator com as variáveis: Inverso do Incremento do Desmatamento, Rebanho Bovino e Nº de Estabelecimentos Agrícolas. Já o terceiro fator é composto por PIB e ICMS.

**Análise Descritiva dos Dados**

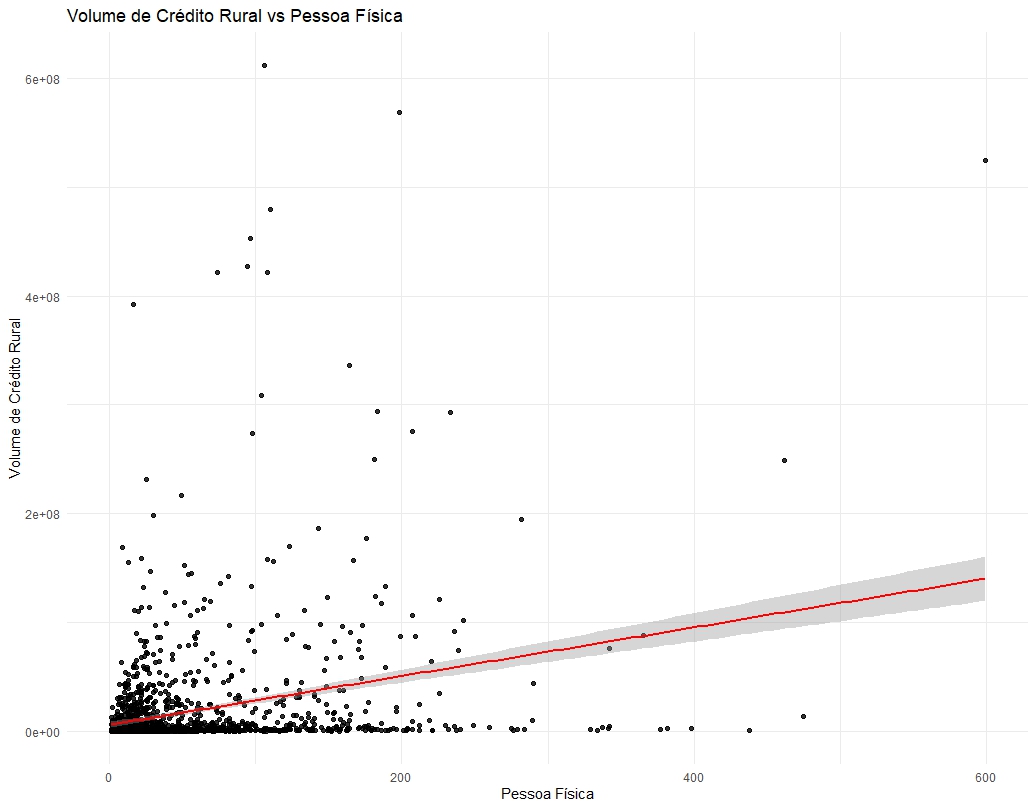




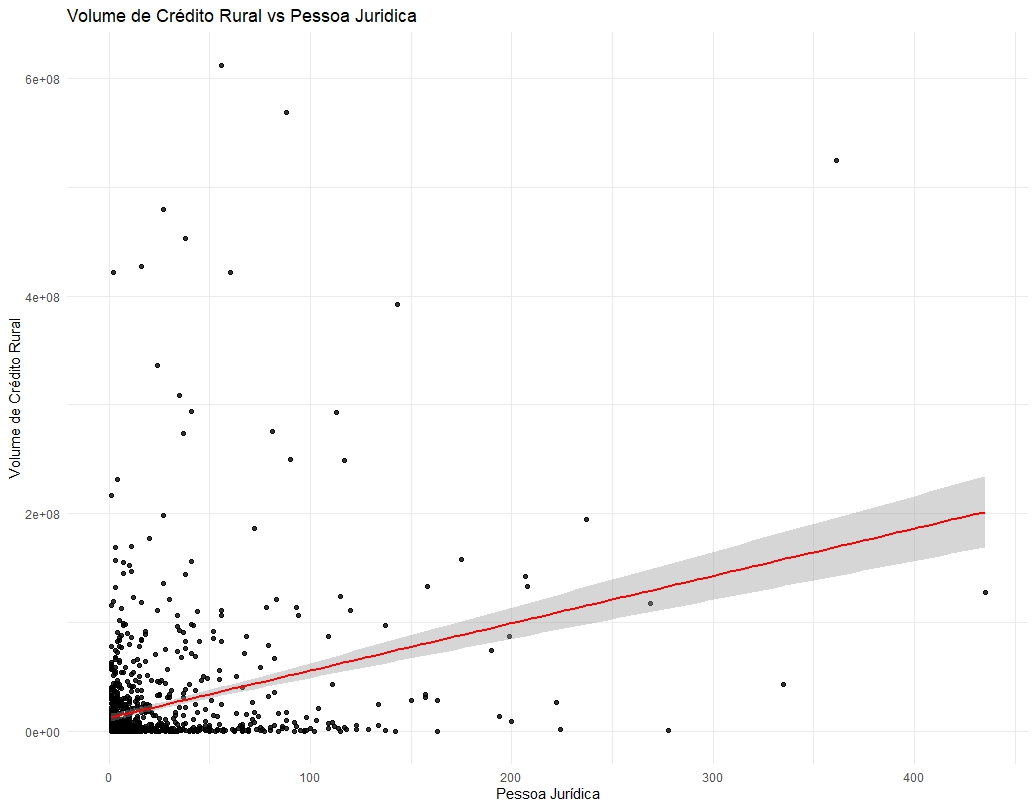
**Matriz de Correlação**



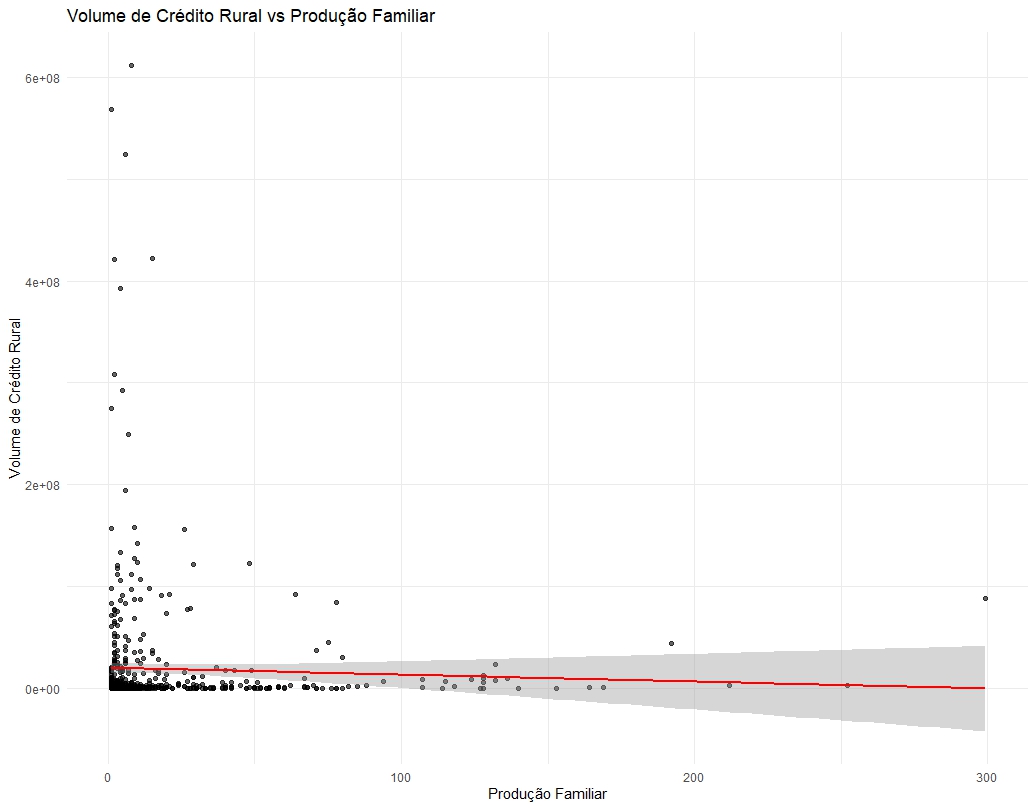
**Gráfico de Dispersão: Volume de Crédito Rural x Pessoa Física**

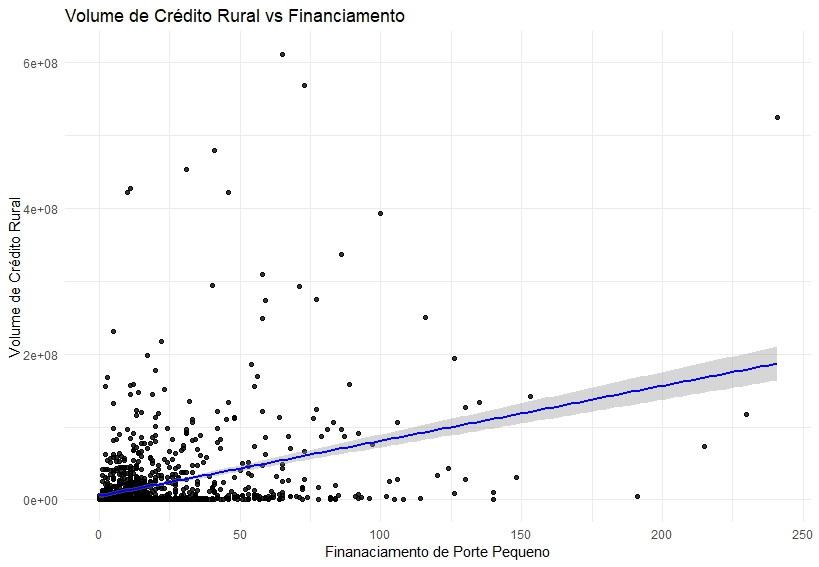


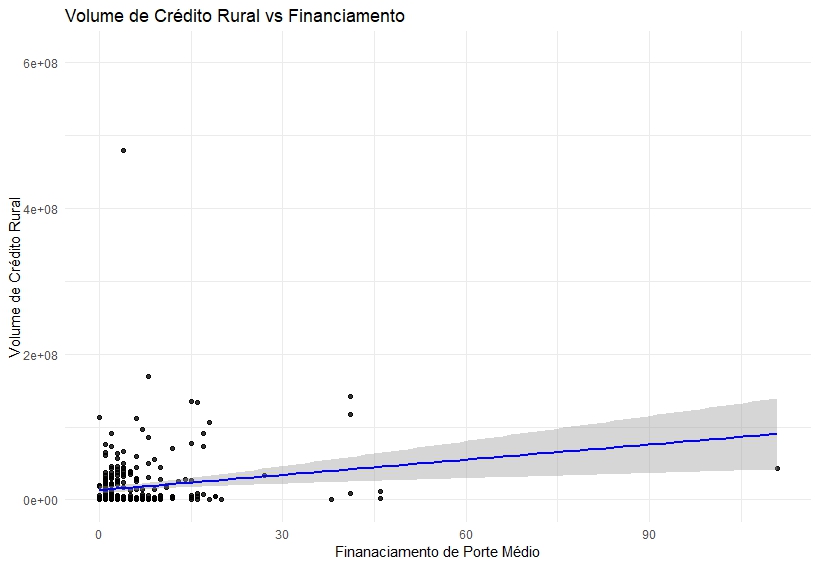
**Gráfico de Dispersão: Volume de Crédito Rural x Pessoa Jurídica**

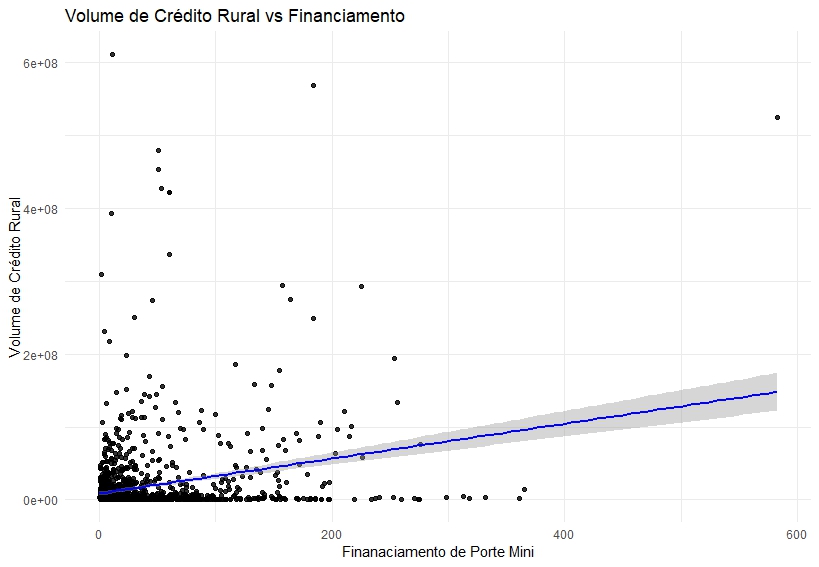


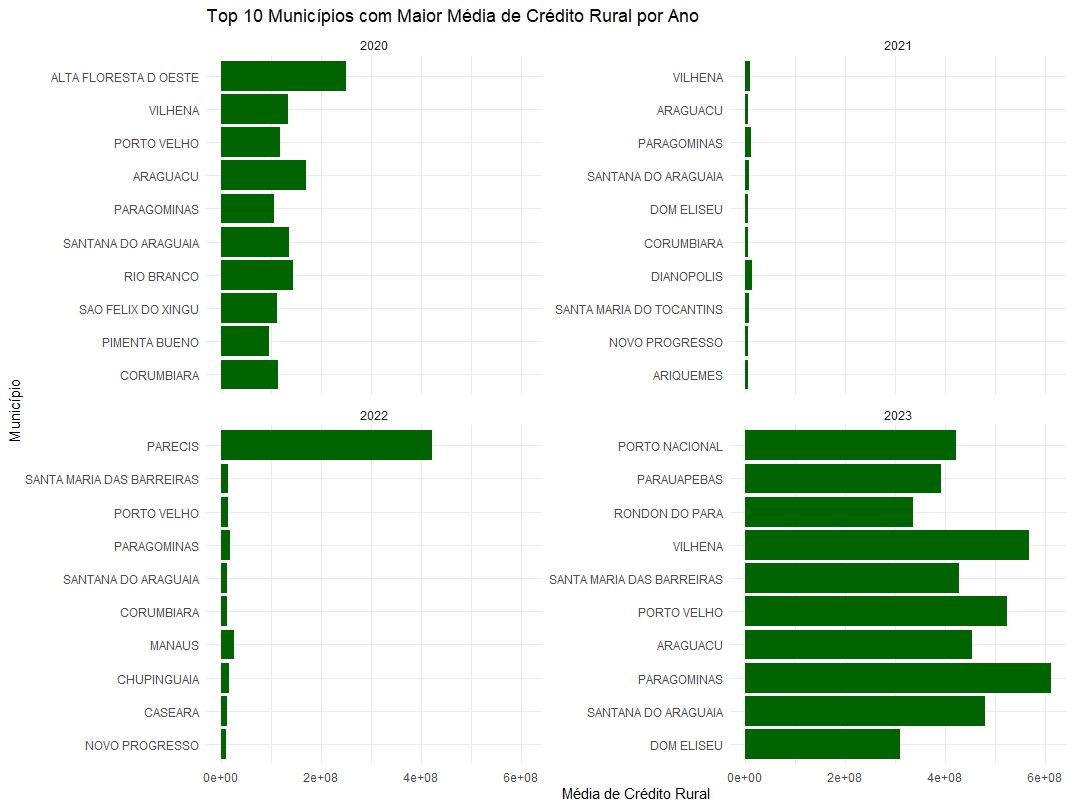
**Gráfico de Dispersão: Volume de Crédito Rural x Produção Familiar**

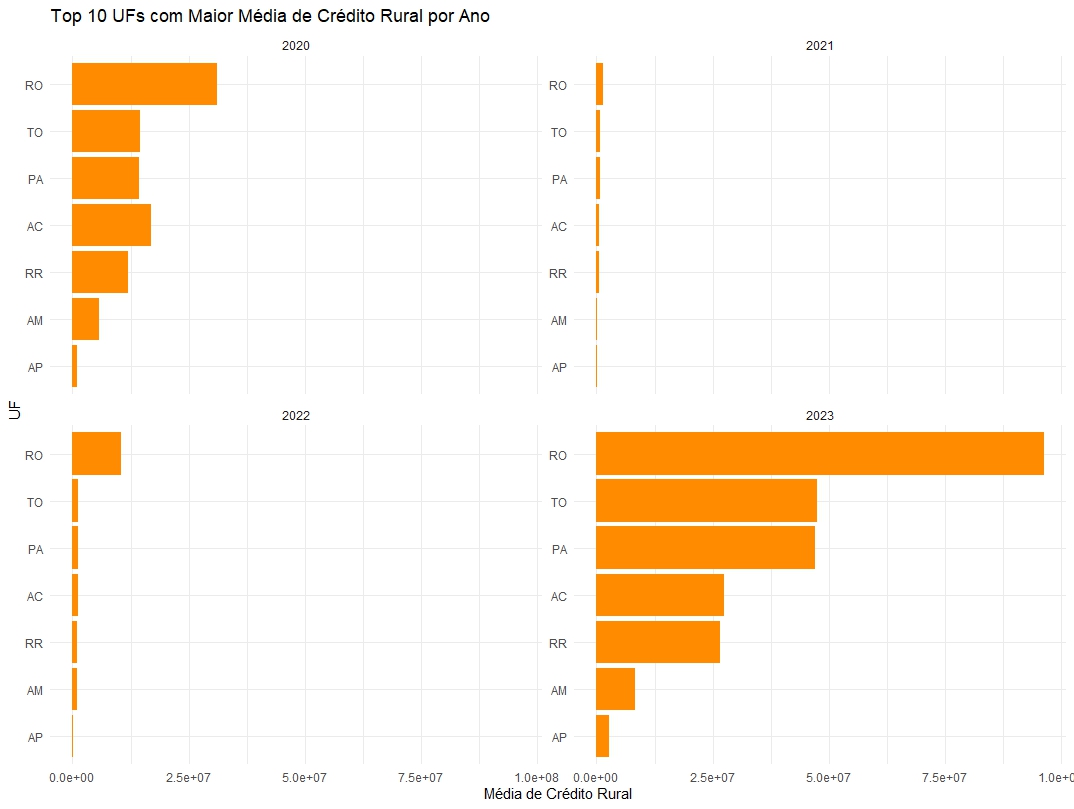


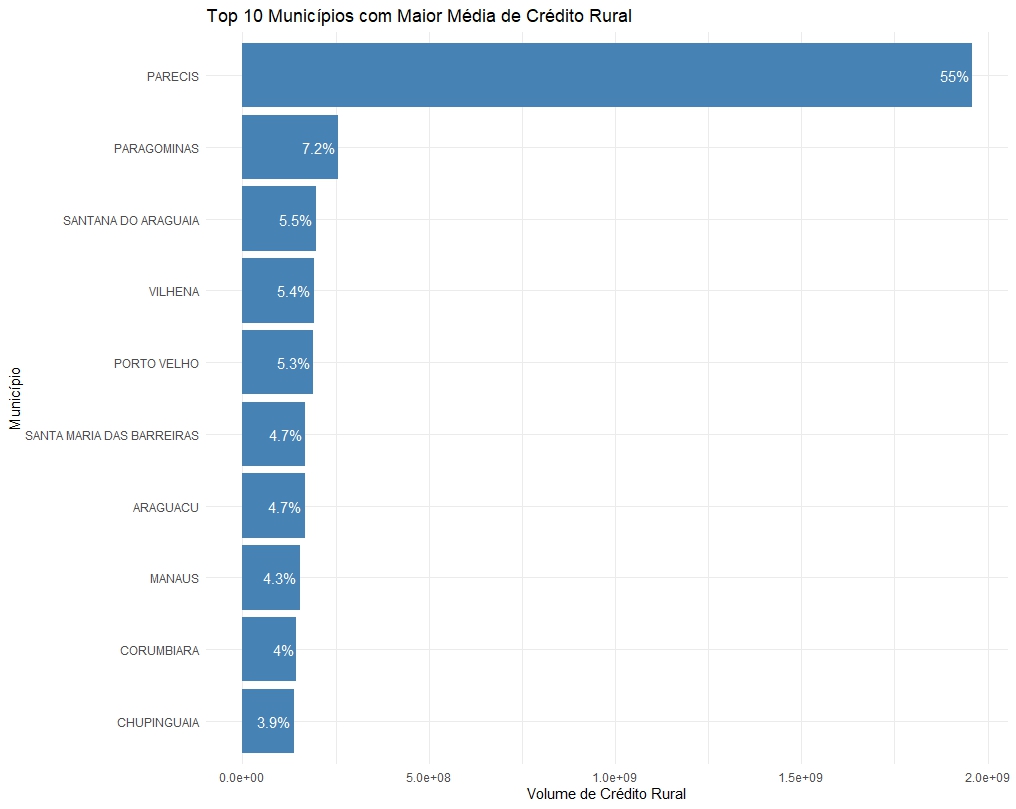


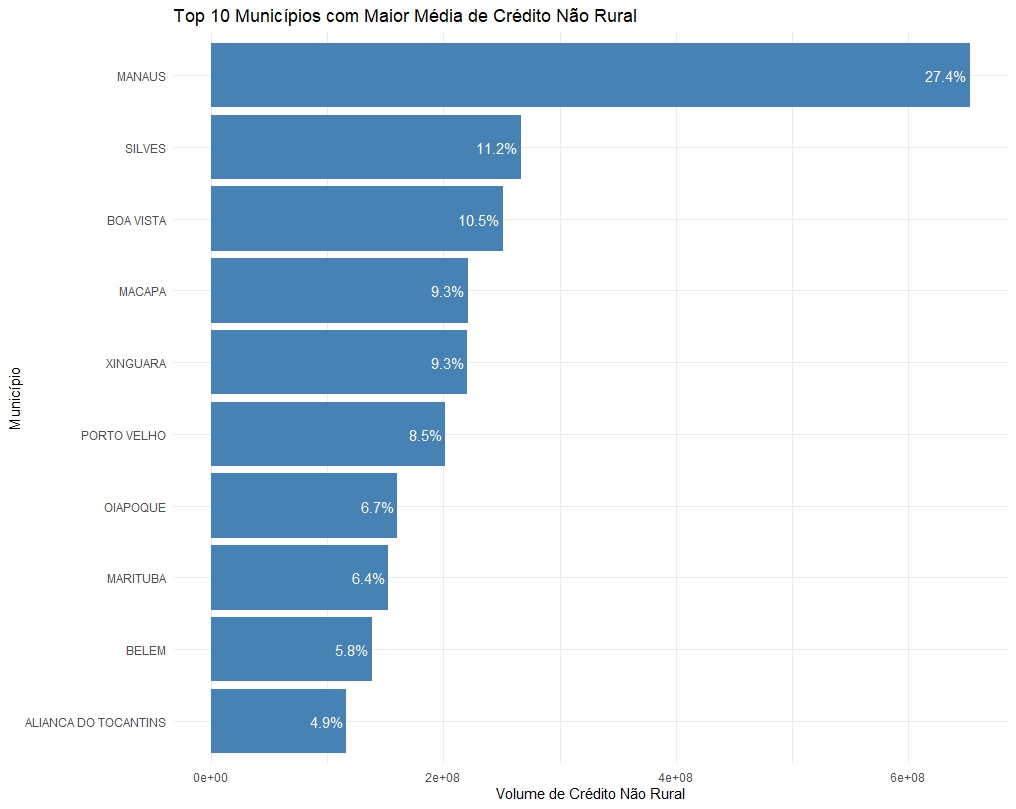












1. **Modelo de Regressão de Dados em Painel**

A análise de dados em painel, também conhecidos como dados longitudinais ou seccionais temporais, envolve observações de múltiplas unidades (como indivíduos, empresas ou em nosso caso “**Municípios**”) ao longo do tempo. A principal vantagem desse tipo de dado é que ele permite controlar tanto a heterogeneidade não observada entre unidades quanto os efeitos da evolução temporal, oferecendo inferências estatísticas mais robustas.

Formalmente, um painel de dados pode ser representado como:

Onde:

* é a variável dependente para a unidade i no tempo t;
* **xit ​** representa uma ou mais variáveis independentes;
* **α** é o intercepto;
* **β** são os coeficientes dos regressores;
* **uit​** é o erro aleatório.

No entanto, diferentes estratégias de modelagem são utilizadas dependendo de como tratamos os efeitos não observados uit​, que podem ser decompostos como:

**Uit = μi + λt + εi**

* μi​: efeito específico da unidade (ex.: município);
* λt ​: efeito específico do tempo (ex.: ano);
* εit​: erro idiossincrático.

A análise de dados em painel combina informações de corte transversal e séries temporais, permitindo explorar variações entre indivíduos ao longo do tempo. Este tutorial cobre os principais modelos para análise de dados em painel, suas especificidades, como escolher o modelo mais adequado, e técnicas para validar as estimativas.

Os modelos abordados incluem:

* **Modelo Pooled OLS (dados empilhados)**: Assume que não há diferenças entre indivíduos.
* **Modelo de Efeitos Fixos (FE)**: Captura variações intra-individuais, controlando para características invariantes.
* **Modelo de Efeitos Aleatórios (RE)**: Assume que as diferenças entre indivíduos são parte do termo de erro.

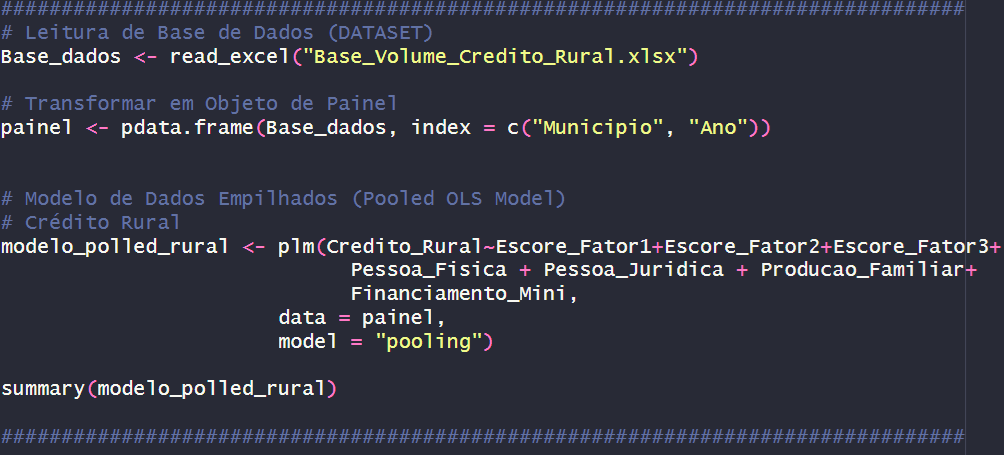
1. **Estimação dos Modelos**
   1. **Modelo de Dados Empilhados (Pooled OLS)**

Este modelo trata de “empilhar” todas as observações da base de dados, ignorando a estrutura de dados em painel. Desta forma, todas as observações são tratadas como não correlacionadas para os indivíduos, com erros homocedásticos para com os indivíduos.

Trata-se, portanto, da forma mais simplista e ingênua pois desconsidera as dimensões de tempo e espaço combinados, ao mesmo tempo que estima a regressão pelo método dos Mínimos Quadrados Ordinários (MQO) (Gujarati e Porter 2011).

Para executar este modelo de regressão é necessário utilizar a função chamada plm do pacote ***plm*** (Croissant e Millo, 2008), juntamente com a variáveis dependente (no caso o Volume de Crédito Rural) e independentes, indicando a base de dados (*data*) e o tipo do modelo de Regressão (“*pooling*”).

**Figura 1**. Código criado na linguagem de Programação R4.4 para realizar o modelo de dados empilhados (Pooled Model).

****

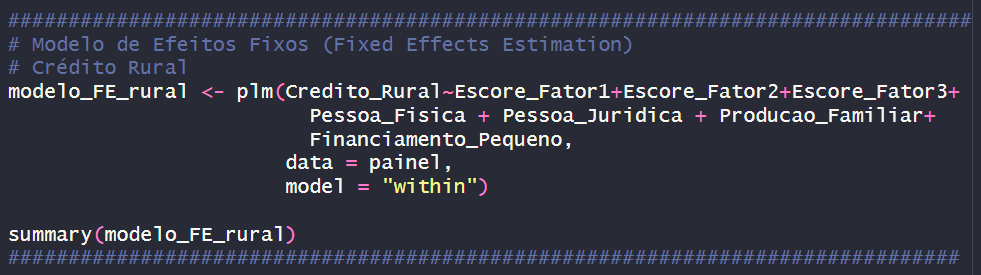
**Tabela 1.** Modelo de Regressão em **Pooled** para dos referentes ao Volume de Crédito Rural.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coeficientes** | **Estimate** | **Std.Error** | **T-value** | **P-value** |
| (Intercept) | 1.8527e+07 | 3.8484e+06 | 4.8143 | 2.06e-06 \*\*\* |
| Escore do Fator 1 | -4.7168e+01 | 2.5476e+01 | -1.8515 | 0.064799 . |
| Escore do Fator 2 | 6.6936e+01 | 2.0702e+01 | 3.2333 | 0.001319 \*\* |
| Escore do Fator 3 | 1.8573e+03 | 1.3273e+01 | 139.9332 | < 2.2e-16 \*\*\* |
| Pessoa Física | 1.4694e+05 | 1.0322e+05 | 1.4236 | 0.155312 |
| Pessoa Jurídica | 1.1701e+05 | 7.4257e+04 | 1.5758 | 0.115815 |
| Produção Familiar | -1.7253e+05 | 1.2343e+05 | -1.3977 | 0.162921 |
| Financiamento de Porte Mini | -1.4408e+0 | 1.1644e+05 | -1.2374 | 0.216628 |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | | | |
| **Total Sum of Squares**: 5.9969e+19 | | | | |
| **Residual Sum of Squares**: 1.249e+18 | | | | |
| **R-Squared**: 0.97917 | | | | |
| **Adj. R-Squared**: 0.97883 | | | | |
| **F-statistic**: 2841.05 on 7 and 423 DF, p-value: < 2.22e-16 | | | | |

* 1. **Modelo de Efeitos Fixos (FE)**

O modelo de efeitos fixos é aquele utilizado para analisar o impacto de variáveis ao longo do tempo. Esse modelo considera os valores de interceptos de cada regressão que variam conforme os efeitos de cada companhia e os coeficientes angulares das variáveis independentes em cada equação que são os mesmos para cada Município do estudo. Desta forma, o intercepto da equação difere de Município para Município, mas os efeitos das variáveis independentes são os mesmo sobre a variável dependente.

**Figura 2**. Código criado na linguagem de Programação R4.4 para realizar o modelo de Efeito Fixos (Fixed Effects Model).

****

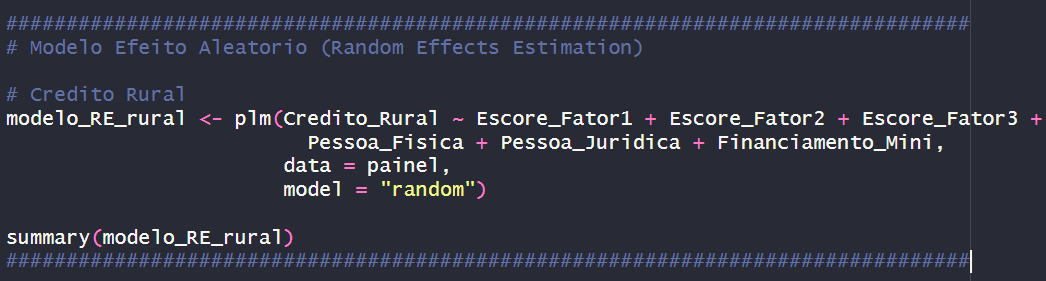
**Tabela 2.** Modelo de Regressão em **Efeitos Fixos** para dados referentes ao Volume de Crédito Rural.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Predictors** | **Estimate** | **Std.Error** | **T-value** | **P-value** |
| Escore do Fator 1 | -2.0597e+02 | 3.5805e+01 | -5.7512 | 2.953e-08 \*\*\* |
| Escore do Fator 2 | -3.1657e+02 | 2.4772e+01 | -12.7793 | < 2.2e-16 \*\*\* |
| Escore do Fator 3 | 1.8857e+03 | 3.8194e+00 | 493.7081 | < 2.2e-16 \*\*\* |
| Pessoa Jurídica | 3.0513e+05 | 6.2581e+04 | 4.8758 | 2.087e-06 \*\*\* |
| Produção\_Familiar | -1.5095e+04 | 2.4608e+04 | -0.6134 | 0.540256 |
| Financiamento Pequeno | -3.2861e+05 | 1.1910e+05 | -2.7591 | 0.006289 \*\* |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | | | |
| **Total Sum of Squares**: 4.9349e+12 | | | | |
| **Residual Sum of Squares**: 1.8454e+11 | | | | |
| **R-Squared**: 0.96 | | | | |
| **Adj. R-Squared**: 0.95 | | | | |
| **F-statistic**: 12479.8 on 6 and 219 DF, p-value: < 2.22e-16 | | | | |

* 1. **Modelo de Efeitos Aleatórios (RE)**

O modelo de efeitos aleatórios, é aquele com efeitos aleatórios onde incluem variações dentro e entre as entidades. No modelo de regressão com efeitos aleatórios, os efeitos individuais dos Municípios são considerados variáveis aleatórias, ao contrário do modelo de efeitos fixos. Esse modelo estatístico apresenta alguma forma de variação aleatória.

**Figura 3**. Código criado na linguagem de Programação R4.4 para realizar o modelo de Efeito Aleatório (Randon Model).



**Tabela 3.** Modelo de Regressão em **Efeitos Aleatórios** para dados referentes ao Volume de Crédito Rural.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Coeficientes** | **Estimate** | **Std.Error** | **Z-value** | **P-value** |
| **(Intercept)** | 1.1929e+07 | 2.3082e+06 | 5.1683 | 2.363e-07 \*\*\* |
| Escore do Fator 1 | -2.3589e+01 | 1.9782e+01 | -1.1924 | 0.233104 |
| Escore do Fator 2 | 1.5936e+02 | 1.5780e+01 | -10.0992 | < 2.2e-16 \*\*\* |
| Escore Fator 3 | 1.8728e+03 | 6.0527e+00 | 309.4055 | < 2.2e-16 \*\*\* |
| Pessoa Física | 2.4214e+05 | 3.8795e+04 | 6.2415 | 4.334e-10 \*\*\* |
| Pessoa Jurídica | 1.1255e+05 | 3.6976e+04 | 3.0438 | 0.002336 \*\* |
| Financiamento de Porte Mini | -9.1085e+04 | 4.1739e+04 | -2.1822 | 0.029093 \* |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | | | |
| **Total Sum of Squares**: 4.6052e+19 | | | | |
| **Residual Sum of Squares**: 5.334e+17 | | | | |
| **R-Squared**: 0.98842 | | | | |
| **Adj. R-Squared**: 0.98836 | | | | |
| **Chisq:** 96210.4 on 6 DF, p-value: < 2.22e-16 | | | | |

**Comparação entre os Modelos**

Após a evidenciação dos modelos de regressão dos tipos agrupados (Pooled), de Efeitos Fixos e de Efeitos Aleatórios, é preciso efetuar os testes para definir qual é o melhor modelo e que por consequência deverá ser considerado.

**Modelo Pooled x Modelo de Efeitos Fixos**

O teste F (F Test for individual Effects) é em homenagem a Gregory Chow. Este teste estatístico é realizado para comprovação da “quebra” de uma tendência estável de uma série histórica e é amplamente utilizada em modelos econométricos.

Atualmente, é usado para verificar se háefeitos individuais significativos no modelo de dados em painel. Esse teste avalia se as diferenças entre os indivíduos (ou unidades) são estatisticamente significativas.

Os argumentos desta função são dois modelos: o primeiro sendo um modelo interno de efeitos fixos e o segundo um modelo de pool. Os efeitos testados são individuais, temporais ou de dois modos, dependendo dos efeitos introduzidos no modelo interno (EF).

**Teste de Hipóteses.**

**H0:** há igualdade nos interceptos e nas inclinações para todos os indivíduos, caracterizando um modelo (Pooled).

**H1**: O modelo de e Efeito Fixo é o mais indicado.

Para calcular o teste usa a função ***pFtest*** () do pacote plm na linguagem R.

**Tabela 4.** Teste F de Chow para dados referentes ao Volume de Crédito Rural.

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **F-Test value** | **P-value** |
| Escore do Fator 1 | 11,53 | 2.2e-16\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 206 | | |
| **df2:** 217 | | |

No teste acima, rejeita-se H0, ou seja, devemos considerar a presença de efeitos específicos. Há **evidências estatísticas fortes** de que os efeitos individuais são significativos. Isso significa que as diferenças entre os indivíduos (ou municípios) têm impacto na variável dependente (Crédito Rural) e, portanto, um modelo que considere esses efeitos (como o modelo de efeitos fixos) é mais apropriado do que um modelo Pooled OLS, que não leva em conta essas diferenças.

**Modelo Pooled x Modelo de Efeitos Aleatórios**

Para comparar os modelos foi realizado o **Teste de Lagrange Multiplier (LM)** ou **Teste de Breusch-Pagan**. Esse teste foi desenvolvido em 1979 por Trevor Breusch e Adrian Pagan, e é utilizado para testar a homoscedasticidade em um modelo de regressão linear, ele avalia a variância dos erros de uma regressão e se os valores dependem das variáveis independentes.

Esse teste foi utilizado para decidir entre um Modelo de regressão OLS agrupada (Pooled) ou uma Regressão de Efeitos Aleatórios.

**Teste de Hipóteses.**

**H0:** Não há variância significativa nos efeitos aleatórios (modelo Pooled OLS é apropriado).

**H1**: Existe variância significativa nos efeitos aleatórios (modelo de efeitos aleatórios é mais apropriado).

Para calcular o teste usa a função ***plmtest*** () do pacote plm na linguagem R.

**Tabela 4.** Teste de Breusch-Paganpara dados referentes ao Volume de Crédito Rural.

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **Chisq-Test value** | **P-value** |
| Escore do Fator 1 | 482.49 | 2.32e-16\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 1 | | |

Isso significa que há evidências de que a variância dos efeitos aleatórios é significativa, o que justifica o uso do modelo de **efeitos aleatórios (RE)** ao invés do modelo Pooled OLS.

**Modelo de Efeitos Fixos x Modelo de Efeitos Aleatórios**

O **Teste de Hausman** é utilizado para decidir entre os modelos de **efeitos fixos (FE)** e **efeitos aleatórios (RE)** em uma análise de dados em painel. Ele verifica se as estimativas dos coeficientes do modelo de efeitos aleatórios são consistentes em relação ao modelo de efeitos fixos.

**Teste de Hipóteses.**

**H0:** O modelo de efeitos aleatórios (RE) é consistente e, portanto, apropriado.

**H1**: O modelo de efeitos aleatórios (RE)é inconsistente, indicando que o modelo de efeitos fixos (FE) é mais apropriado.

Para calcular o teste usa a função ***phtest*** () do pacote plm na linguagem R.

**Tabela 4.** Teste de Hausmanpara dados referentes ao Volume de Crédito Rural.

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **Chisq-Test value** | **P-value** |
| Escore do Fator 1 | 23,67 | 0,00025\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 5 | | |

O p-valor é muito baixo (menor que 0,05), levando a rejeição da hipótese nula. Isso significa que o modelo de **efeitos fixos** é consistente e apropriado para os dados.

**Pressupostos do Modelo**

1. **Teste de Correlação Cross-Section ou Contemporânea**

Para investigar se existe a dependência cross-section, é aplicado o de Breusch-Pagan LM.

**Teste de Hipóteses.**

**H0:** Não tem correlação cross-section

**H1**: Tem Correlação

Para calcular o teste usa a função ***pcdtest*** (*modelo, test = c ("lm")*) do pacote plm na linguagem R.

**Tabela 4.** Teste de Breusch-Pagan LM

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **Chisq-Test value** | **P-value** |
| Escore do Fator 1 | 9754,9 | 2.2e-16\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 5439 | | |

O p-valor é muito baixo (menor que 0,05), levando a rejeição da hipótese nula. Isso significa que o modelo de **efeitos fixos** possui dependência cross-section. Assim, é preciso corrigir erros robustos.

1. **Teste de Correlação Serial**

Para investigar se existe a dependência serial, é aplicado o de Breusch-Godfrey/Wooldrige.

**Teste de Hipóteses.**

**H0:** Não tem correlação serial nos erros idiossincráticos

**H1**: Tem Correlação serial nos erros idiossincráticos

Para calcular o teste usa a função ***pbgtest*** () do pacote plm na linguagem R.

**Tabela 4.** Teste de Breusch-Godfrey/Wooldrige

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **Chisq-Test value** | **P-value** |
| Escore do Fator 1 | 25,02 | 5,72e-07\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 1 | | |

O p-valor é muito baixo (menor que 0,05), levando a rejeição da hipótese nula. Isso significa que o modelo de **efeitos fixos** possui correlação serial nos erros idiossincráticos. Com isso, é preciso corrigir erros robustos.

1. **Teste de Heterocedasticidade**

**Teste de Hipóteses.**

**H0:** homocedásticos

**H1**: heterocedasticos

Para calcular o teste usa a função ***bptest*** () do pacote lmtest na linguagem R.

**Tabela 4.** Teste de Breusch-Pagan

|  |  |  |
| --- | --- | --- |
| **Coeficientes** | **Chisq-Test value** | **P-value** |
| Escore do Fator 1 | 1809,3 | 2.2e-16\*\*\* |
| Escore do Fator 2 |
| Escore do Fator 3 |
| Pessoa Física |
| Pessoa Jurídica |
| Financiamento de Porte Mini |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | |
| **df1**: 7 | | |

O p-valor é muito baixo (menor que 0,05), levando a rejeição da hipótese nula. Isso significa que o modelo de **efeitos fixos** possui erros heterocedasticos.

**MODELO DE REGRESSÃO DE DADO EM PAINEL FINAL**

Neste estudo, foi aplicada a metodologia de Driscoll e Kraay (1998), que propõe um estimador robusto para a matriz de variância-covariância, corrigindo para possíveis Heterocedasticidade, autocorrelação e dependência cruzada entre os municípios. Essa abordagem foi utilizada para ajustar o modelo selecionado com base em testes prévios, resultando em uma versão final que melhor explica o volume de crédito rural.

**Tabela 2.** Modelo de Regressão em **Efeitos Fixos** **Corrigido** para dados referentes ao Volume de Crédito Rural.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Predictors** | **Estimate** | **Std.Error** | **T-value** | **P-value** |
| Escore do Fator 1 | -1.7920e+02 | 1.5094e+02 | -1.1872 | 0.2364348 |
| Escore do Fator 2 | -3.0559e+02 | 4.3974e+0 | -6.9494 | 4.226e-11 \*\*\* |
| Escore do Fator 3 | 1.8849e+03 | 6.6173e+00 | 284.8460 | < 2.2e-16 \*\*\* |
| Pessoa Física | 1.2129e+05 | 3.4449e+04 | 3.5209 | 0.0005241 \*\*\* |
| Pessoa Jurídica | 2.3927e+05 | 8.5012e+04 | 2.8145 | 0.0005241 \*\*\* |
| Produção\_Familiar | -9.3545e+04 | 1.0471e+05 | -0.8933 | 0.3726604 |
| Financiamento Pequeno | -2.6733e+05 | 9.9798e+04 | -2.6787 | 0.0079568 \*\* |
| **Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1** | | | | |
| **Total Sum of Squares**: 3.9349e+19 | | | | |
| **Residual Sum of Squares**: 1.0454e+17 | | | | |
| **R-Squared**: 0.997 | | | | |
| **Adj. R-Squared**: 0.995 | | | | |
| **F-statistic**: 11637.8 on 7 and 217 DF, p-value: < 2.22e-16 | | | | |

O modelo de regressão com dados em painel e efeitos fixos foi estimado para investigar os determinantes do crédito rural concedido aos municípios ao longo do tempo. A estrutura em painel permitiu controlar a heterogeneidade não observada entre os municípios, isolando os efeitos das variáveis explicativas incluídas no modelo.

Os resultados mostram que o modelo possui alto poder explicativo, com um R² ajustado de 99,47%, indicando que quase toda a variação no volume de crédito rural entre os municípios é explicada pelas variáveis selecionadas. O teste de Hausman confirmou a superioridade do modelo de efeitos fixos sobre o de efeitos aleatórios, justificando a escolha metodológica.

Entre os fatores utilizados no modelo foram extraídos por meio da análise fatorial, agrupando variáveis com alta correlação entre si, Entre eles, destacam-se:

* **Fator 3 (PIB e ICMS)** apresentou um coeficiente positivo e altamente significativo, indicando que municípios com maior dinamismo econômico atraem volumes superiores de crédito rural. Esse resultado é consistente com a hipótese de que áreas com maior geração de riqueza e arrecadação têm melhores condições de acesso ao sistema de crédito, seja por infraestrutura financeira mais consolidada ou pela maior capacidade de oferecer garantias.
* Em contrapartida, o **Fator 2 (Rebanho bovino, Nº de Estabelecimentos Agrícolas e Inverso do Incremento do Desmatamento)** teve efeito negativo e também altamente significativo. Este resultado pode refletir um perfil de municípios mais rurais e conservadores ambientalmente, que, apesar de relevantes para o setor agropecuário, não concentram os maiores volumes de crédito. Isso pode apontar para uma concentração de recursos em regiões com perfil mais empresarial ou intensivo em capital.
* O **Fator 1 (População, empregos formais, educação básica e agências bancárias)** não apresentou significância estatística, sugerindo que, uma vez controlado o efeito fixo dos municípios, essas dimensões urbanas e sociais não influenciam diretamente o volume de crédito rural concedido.

O coeficiente positivo do **Escore do Fator 3** indica que cada aumento de um desvio padrão na combinação de PIB e ICMS de um município está associado a um acréscimo médio de aproximadamente **R$ 1.884.000,90** no volume de crédito rural concedido. Esse resultado reforça a importância da capacidade econômica local no acesso ao crédito rural, evidenciando que municípios com maior atividade econômica tendem a receber mais recursos financeiros destinados ao setor.

Por outro lado, o coeficiente negativo do **Escore do Fator 2** revela que cada aumento de um desvio padrão na combinação de Rebanho Bovino, Número de Estabelecimentos Agrícolas e Inverso do Incremento do Desmatamento está associado a uma redução média de aproximadamente **R$ 305.590,00** no volume de crédito rural concedido. Esse resultado sugere que municípios com características agropecuárias mais consolidadas ou com menor pressão recente por desmatamento podem estar recebendo menos crédito novo, possivelmente em função de menor demanda por expansão produtiva ou por não estarem entre as prioridades das instituições financeiras.

Em relação às variáveis que representam o tipo de tomador, o coeficiente positivo da variável **Pessoa Física** indica que cada aumento de R$ 1,00 no volume de crédito rural destinado a pessoas físicas está associado a um aumento médio de aproximadamente **R$ 121.290,00** no volume total de crédito rural no município, mantidas constantes as demais variáveis. Esse resultado é estatisticamente significativo (p < 0,001), reforçando o papel das operações de crédito individual como um dos principais determinantes do volume de crédito rural concedido localmente.

De forma similar, o coeficiente positivo da variável **Pessoa Jurídica** indica que cada aumento de R$ 1,00 no volume de crédito rural destinado a pessoas jurídicas está associado a um aumento médio de aproximadamente **R$ 239.270,00** no volume total de crédito rural no município. Esse efeito também é estatisticamente significativo (p < 0,01), sugerindo que as operações de crédito rural voltadas a empresas e produtores formais exercem papel substancial na composição do crédito total concedido.

O coeficiente da variável **Produção Familiar** foi negativo (**R$ -93.545,00**), porém não estatisticamente significativo (**p = 0,372**). Isso indica que o volume de crédito rural destinado à agricultura familiar não apresentou relação estatisticamente robusta com o volume total de crédito rural nos municípios, dentro do modelo ajustado. Embora seja um resultado inconclusivo, ele pode refletir limitações no acesso ao crédito por esse segmento ou variações que não foram capturadas pelas demais variáveis controladas.

Por sua vez, a variável **Financiamento Pequeno** apresentou um coeficiente negativo e estatisticamente significativo (**R$ -267.330,00; p = 0,008**). Isso sugere que à medida que aumenta o volume de operações classificadas como de pequeno porte, há uma tendência de redução no volume total de crédito rural concedido. Esse resultado pode indicar que, embora as operações pequenas sejam mais numerosas, seu impacto agregado no montante total de crédito é limitado, o que pode revelar uma concentração de valores em grandes operações de crédito.

Os resultados do modelo indicam que o **acesso ao crédito rural** está fortemente relacionado à **estrutura econômica local**, conforme demonstrado pelo impacto positivo e significativo do **PIB e ICMS (Fator 3)**. Isso evidencia que municípios com maior dinamismo econômico tendem a atrair mais recursos financeiros, possivelmente devido a maior capacidade de pagamento, infraestrutura ou atratividade para instituições financeiras.

Por outro lado, o efeito negativo do **Fator 2**, que representa características agropecuárias consolidadas e menor pressão de desmatamento, pode refletir uma tendência de **deslocamento do crédito para áreas em expansão agrícola ou em transformação econômica**, em detrimento de áreas já consolidadas.

Além disso, os resultados mostram que **o crédito rural destinado a pessoas físicas e jurídicas** é um importante determinante do volume total de crédito, com destaque para o papel expressivo das operações com pessoas jurídicas, que representam maior valor agregado. No entanto, o **efeito não significativo da produção familiar** e o **efeito negativo do financiamento de pequeno porte** apontam para uma possível desigualdade no acesso ao crédito por parte dos pequenos produtores e da agricultura familiar — justamente os segmentos que muitas vezes mais dependem de políticas públicas específicas para viabilizar sua produção.

Nesse sentido, os resultados do estudo podem subsidiar **políticas de crédito rural mais equitativas**, com **foco na ampliação do acesso ao crédito para pequenos produtores** e maior valorização da agricultura familiar. Além disso, reforçam a importância de estratégias que considerem o perfil socioeconômico dos municípios e promovam uma **distribuição mais equilibrada dos recursos**, respeitando os diferentes estágios de desenvolvimento rural.

**REFERÊNCIAS BIBLIOGRÁFICAS**

CROISSANT, Y; MILO, G. “Econometria de dados em painel em R: o pacote plm.” *Journal of Statistical Software*, *27* (2), 1-43, 2018. doi: 10.18637/jss.v027.i02 (URL: <http://doi.org/10.18637/jss.v027.i02> ).

GREENE, William H. Análise econométrica. 7ª ed. Pearson Education, 2012.

TORRES-REYNA, Oscar. Introdução a Modelos de Efeitos Fixos/Aleatórios usando R. Princeton: Universidade de Princeton, 2010. Disponível em: <http://www.princeton.edu/~otorres/Panel101R.pdf> .

Breusch, Trevor S. 1978. «Testing for autocorrelation in dynamic linear models». Australian Economic Papers 17 (31): 334–55.

Breusch, Trevor S, e Adrian R Pagan. 1979. «A simple test for heteroscedasticity and random coefficient variation». Econometrica: Journal of the Econometric Society, 1287–94.

Breusch, Trevor Stanley, e Adrian Rodney Pagan. 1980. «The Lagrange multiplier test and its applications to model specification in econometrics». The Review of Economic Studies 47 (1): 239–53.

Gujarati, Damodar N., e Down C Porter. 2011. Econometria básica. 5a ed. New York: Mc Graw Hill. <https://doi.org/10.1126/science.1186874>.

Hausman, Jerry A. 1978. «Specification tests in econometrics». Econometrica: Journal of the econometric society, 1251–71.

Pesaran, M Hashem. 2015. «Testing weak cross-sectional dependence in large panels». Econometric Reviews 34 (6-10): 1089–1117.

Wooldrige, Jeffrey M. 2010. Econometric analysis of cross section and panel data. MIT Press.