1. **INTRODUÇÃO**

A análise de regressão ocupa-se do estudo da dependência de uma variável, a variável dependente, em relação a uma ou mais variáveis, as variáveis explicativas, com o objetivo de estimar e/ou prever a média (da população) ou o valor médio da dependente em termos dos valores conhecidos ou fixos (em amostragem repetida) das explicativas (ANDERSON et al., 2007).

1. **MATERIAIS E MÉTODOS**

Os dados foram obtidos no site da... essas informações devem ser fornecidas (ou preenchidas) pelo cliente.

**2.1 FORMULAÇÃO MATEMÁTICA**

**2.1.1 Regressão linear múltipla**

Para Hair, Anderson Taham & Black (2005) a análise de regressão múltipla é uma técnica estatística, que pode ser usada para analisar a relação de causa e efeito entre uma única variável dependente e diversas variáveis independentes.

A análise de regressão múltipla tem por objetivo, conforme Fávero, Belfiore, Silva & Chan (2009), estimar o impacto do incremento de cada variável independente – que se traduz como peso de cada variável independente – sobre a respectiva variação da variável dependente. De acordo com Hair, Anderson Taham & Black (2005), os pesos denotam a contribuição relativa das variáveis independentes para a previsão geral e facilitam a interpretação sobre a influência de cada variável explicativa em fazer a previsão.

O modelo de regressão múltipla é dado por (CHARNET, 2008):

 (1)

Em que: é o fenômeno em estudo (variável dependente);representa o intercepto (constante); são os coeficientes associados a cada variável independente (coeficientes angulares); *xki* são as variáveis explicativas (independentes) com ;  é o termo do erro. O erro , também chamado de resíduo, representa possíveis variáveis que não foram inseridas no modelo, mas que também contribuiriam para a explicação de  (CHARNET, 2008).

**2.1.2 Adequação do modelo**

Após estimar o modelo é necessário verificar se de fato representa o que deseja-se medir, isto é, o modelo expressa a realidade e, consequentemente, proporciona uma melhoria nas decisões (CORRAR *et. al.,* 2007). Portanto, avaliou-se: a qualidade do ajuste por meio dos coeficientes de determinação (*R*²) e determinação ajustado (*R*²*ajustado*) e a multicolinearidade entre as variáveis explicativas por meio da análise dos fatores de inflação da variância (*VIF*).

**2.1.2.1 Multicolinearidade**

Outro aspecto importante no ajuste de modelos de regressão linear múltipla é a multicolinearidade. Objetiva-se investigar se há multicolinearidade entre as variáveis explicativas, visto que a forte correlação entre elas acarreta vários efeitos negativos no ajuste do modelo de regressão. A multicolinearidade é um problema comum em regressão linear múltipla, indicando que existe uma relação de linearidade entre as variáveis regressoras, prejudicando assim a estimação dos coeficientes de regressão. O problema de multicolinearidade torna a estimativa dos parâmetros imprecisa, por conta de um alto valor do erro padrão, o que não é conveniente estatisticamente (KUTNER et al., 2004; TAMHANE, DUNLOP, 2000).

A presença de multicolinearidade pode ser observada quando o coeficiente de determinação  apresenta um alto valor, mas nenhum dos coeficientes da regressão é estatisticamente significativo. Portanto, uma das formas de detecção é avaliar o Fator de Inflação da Variância (*VIF*). Esse fator mede o grau de associação entre as variáveis explicativas, a partir do coeficiente de determinação do modelo de regressão ajustado apenas entre as variáveis independentes. O Fator de Inflação da Variância é definido como (BERK, 1977):



Pode-se observar que, quanto maior o , maior é o valor de *VIF*, indicando alta colinearidade.

Valores de *VIFi* maiores que 10 correspondem a um coeficiente de determinação (KUTNER et al., 2004; TAMHANE & DUNLOP, 2000). Partindo dessa idéia/princípio de que coeficiente de determinação acima de 90% (entre variáveis explicativas) caracteriza presença de multicolinearidade, então, justifica-se o porquê de considerar valores de *VIF*i maiores que 10, pois 

**2.1.3 Seleção de variáveis**

Os métodos de seleção têm sido desenvolvidos para identificar se é necessário a inclusão de todas as variáveis regressoras disponíveis ou incluir apenas um subconjunto destas variáveis. No estudo em questão utilizamos o método “passo atrás” (backward), ele consiste na eliminação de uma variável que exerce grande influência sob a soma de quadrados dos resíduos. O ideal é encontrar critérios que balanceiam as duas ideias seguintes, de tendência oposta: (i) por um lado, para melhorar o processo de previsão, maior número possível de variáveis é desejado, pela melhoria de ajustamento decorrente; (ii) por outro lado, devido aos eventuais custos de obtenção dos *x*i’s e às dificuldades de interpretação e de manuseio de um modelo grande, o número de variáveis deve ser tão pequeno quanto possível (DRAPER & SMITH, 1998; CHARNET, 2008).

**2.1.4 Método “passo atrás” (backward)**

Este procedimento caracteriza-se por incorporar, inicialmente, todas as variáveis auxiliares em um modelo de regressão linear múltipla e percorrer etapas, nas quais uma variável por vez pode vir a ser eliminada. Se em uma dada etapa não houver eliminação de alguma variável, o processo é então interrompido e as variáveis restantes definem o modelo final (CHARNET, 2008).

Como critério de eliminação das variáveis, analisam-se as contribuições individuais das variáveis no modelo. A variável de pior desempenho é eliminada, a não ser que esta atenda a um critério mínimo. No julgamento de uma variável, comparamos o modelo completo com o modelo reduzido, pela retirada de tal variável. Observa-se a soma de quadrados de regressão extra assim como a estatística do teste de sua contribuição, dada por:

 (2.13)

Sendo  e  calculadas sob o modelo completo e  calculada sob o modelo reduzido. A estatística (2.13) testa a contribuição da variável após a inclusão das demais. A contribuição é significante se o valor da estatística for maior que um quantil especificado da distribuição *F* com 1 e (*n - p*) graus de liberdade, sendo *p* o número de parâmetros do modelo completo. Assim, se o valor da estatística for menor que esse quantil da distribuição *F*, a contribuição não é considerada significante e o modelo reduzido deverá ser preferido. Se observarmos várias variáveis não significantes, apenas uma variável é eliminada em uma etapa (aquela cuja estatística do teste tiver o menor valor). Quando uma variável é eliminada, passamos para a nova etapa cujo modelo completo não contém a variável que foi descartada. Se todas as variáveis são significantes, o processo é concluído, e o modelo completo desta etapa é o modelo final (CHARNET, 2008).

1. **APLICAÇÃO**

## **3.1 - Modelos Empíricos**

Definiu-se como Modelo 1 aquele que teria como regressores os critérios estipulados pela Lei 13.803 de 2000 (MINAS GERAIS, 2000) e como Modelo 2 os da Lei 18.030 de 2009 (MINAS GERAIS, 2009). As variáveis dependentes se associaram à esses modelos com base nos exercícios de suas ocorrências, conforme Figura 2:

Figura 02 – Definição dos Modelos de Regressão

IMRS

Lei 13.803/2000

Modelo 1

Modelo 2

Lei 18.030/2009

PIB-M

IRFS

IDH-M

PIB-M

IMRS

IRFS

Fonte: Elaborado pelo autor

Para os devidos testes, totalizaram sete regressões distintas entre as alternâncias das variáveis dependentes e independentes, conforme apresentada na Tabela 3.

**Tabela 3 – Relação de modelos aplicados a cada período e quantitativo de regressões.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Variável dependente** | **Período** | **Modelo estatístico** | **Número de regressões** |
| PIB-M | 2002 a 2010 | Modelo 1 | 1 |
| 2011 a 2013 | Modelo 2 | 1 |
| IRFS | 2003 a 2010 | Modelo 1 | 1 |
| 2011 | Modelo 2 | 1 |
| IMRS | 2002, 2004, 2006, 2008 e 2010 | Modelo 1 | 1 |
| 2012 | Modelo 2 | 1 |
| IDH-M | 2010 | Modelo 1 | 1 |
| **Total** | | | **7** |

Fonte: Elaborado pelo autor

Portanto, estimou-se os modelos da seguinte maneira:

### **3.1.1 - Modelo 1**

*Yt = β0 + β1VAFit + β2AGEit + β3POPit + β4POP50it + β5EDUit + β6PALIit + β7PCULit + β8MAMBit + β9SAUit + β10RECPit + β11COTMit + β12MUNMit + єit*

Onde,

Yt = Variável dependente no tempo t”;

β0 = Termo constante do modelo econométrico;

β1...β12 = Parâmetros das variáveis independentes, que serão estimados pela regressão múltipla e que podem explicar a variação do Y;

VAFt = Valor Adicionado Fiscal no tempo “t”;

AGEt = Área geográfica no tempo “t”;

POPt = População no tempo “t”;

POP50t = População dos 50 municípios mais populosos no tempo “t”;

EDUt = Educação no tempo “t”;

PALIt = Produção de Alimentos no tempo “t”;

PCULt = Patrimônio Cultural no tempo “t”;

MAMBt = Meio ambiente no tempo “t”;

SAUt = Saúde no tempo “t”;

RECPt = Receita Própria no tempo “t”;

COTMt = Cota Mínima no tempo “t”;

MUNM t = Municípios Mineradores no tempo “t”;

єt = O termo que representa o erro da estimação.

### **3.5.2 - Modelo 2**

*Yt = β0 + (variáveis independentes do modelo 1) + β13RHIDt + β14EPENt + β15ESPt + β16TURt + β17ICMSSt + β18MPCt + єt*

Onde,

Yt = Variável dependente no tempo “t”;

β0 = Termo constante do modelo econométrico;

β1...β18 = Parâmetros das variáveis independentes, que serão estimados pela regressão múltipla e que podem explicar a variação do I;

RHIDt = Recursos Hídricosno tempo “t”;

EPENt = Sede de estabelecimentos penitenciário no tempo “t”;

ESPt = Esporte no tempo “t”;

β16TURt = Turismo no tempo “t”;

β17ICMSSt = ICMS solidário no tempo “t”;

β18MPCt = Mínimo Per Capita no tempo “t”;

єt = O termo que representa o erro da estimação.

Uma vez definidos os modelos, bastou apenas substituir o termo Y das equações pelas variáveis dependentes desejadas, adequar os períodos conforme particularidades de publicação dos dados e efetuar as respectivas regressões.

1. **RESULTADOS E DISCUSSÃO**

Para darmos início às discussões, primeiro faz-se necessário uma análise descritiva das nossas variáveis, analisando seus comportamentos e tendências.

* 1. **ANÁLISE GERAL**

**4.1.1 Variáveis Dependentes**

O PIB do Vale do Jequitinhonha, resultado da soma das 51 cidades que o compõe, apresenta um comportamento de constante crescimento. Sua menor evolução, conforme ilustrado no Gráfico 1, aconteceu entre os anos de 2008 e 2009, resultado este que pode ser explicado pelas eleições deste mesmo ano e pela situação econômica que nosso país se ingressava.

Gráfico 1 – Comportamento da variável dependente “PIB” no intervalo estudado.

Já o IRFS e o IMRS, representadas pela média das 51 cidades que o compões, apresentaram comportamentos crescentes até o ano de 2008, no qual ambos tiveram uma queda significativa, fato este que pode ser explicado pelos mesmos motivos previamente ditos.

Gráfico 2 – Comportamento da variável dependente “IRFS” no intervalo estudado.

Gráfico 3 – Comportamento da variável dependente “IMRS” no intervalo estudado.

Um dos fatores de extrema importância para a construção de um modelo de regressão linear é a frequência com que as variáveis se apresentam. No Modelo 1 percebemos que a variável IDH-M possui a menor frequência, fato este que pode comprometer a integridade do modelo.

Tabela 4 – Comportamento das variáveis dependentes para o Modelo 1.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variável** | **Frequência** | **Mínimo** | **Máximo** | **Média** | **Desvio Padrão** |
| PIB-M | 459 | 6340 | 355355 | 52307,73 | 51106,338 |
| IRFS | 356 | 0,3870 | 0,5910 | 0,481264 | 0,0358494 |
| IMRS | 255 | 0,3920 | 0,6380 | 0,533373 | 0,0508230 |
| IDH-M | 51 | 0,5410 | 0,7160 | 0,616078 | 0,0334244 |

O mesmo acontece na construção do Modelo 2, visto que a frequência para análise das variáveis dependentes IRFS e IMRS são consideravelmente baixas.

Tabela 5 – Comportamento das variáveis dependentes para o Modelo 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variável** | **Frequência** | **Mínimo** | **Máximo** | **Média** | **Desvio Padrão** |
| PIB-M | 153 | 16712 | 563264 | 100824,69 | 97923,678 |
| IRFS | 43 | 0,4370 | 0,6030 | 0,503860 | 0,0341109 |
| IMRS | 51 | 0,4690 | 0,6040 | 0,542980 | 0,0357500 |

* + 1. **Variáveis Independentes**

Totalizaram 14 variáveis no Modelo 1, na qual 2 delas foram excluídas do modelo por apresentar casos insuficientes, sendo elas “População dos 50 mais populosos” e “Mateus Leme/Mesquita”, conforme mostra na tabela 6.

Tabela 6 – Comportamento das variáveis independentes para o Modelo 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variável** | **Frequência** | **Mínimo** | **Máximo** | **Média** | **Desvio Padrão** |
| VAF | 459 | 6959,1400 | 41826467,8800 | 877815,231351 | 4119868,9991400 |
| Área Geográfica | 459 | 4366,5000 | 367074,3200 | 62503,991503 | 56205,2936505 |
| População | 459 | 9199,3800 | 1617257,6000 | 92051,836427 | 158267,7784331 |
| População dos 50 mais populosos\* | 9 | 849586,0800 | 2273728,0000 | 1523410,796667 | 478635,9083325 |
| Educação | 394 | 335,5500 | 454418,5900 | 153661,560406 | 82666,7578599 |
| Produção de Alimentos | 459 | 3949,3600 | 205439,9500 | 45678,756122 | 30231,4851659 |
| Mateus Leme/Mesquita\* | 21 | 0,0100 | 49209,4000 | 2550,993333 | 10714,6326444 |
| Patrimônio Cultural | 455 | 166,0200 | 557805,2600 | 84511,712088 | 69074,4270245 |
| Meio Ambiente (A+B+C) | 190 | 11,1800 | 542809,2200 | 66564,274684 | 82931,2420000 |
| Programa Saúde da Família | 437 | 56,0500 | 219194,5200 | 34823,597483 | 29624,3501463 |
| Saúde per capita | 459 | 124,2500 | 95355,3200 | 29606,689237 | 14995,9717613 |
| Receita Própria | 459 | 3748,3900 | 437041,5300 | 56559,816754 | 60103,0835605 |
| Cota Mínima | 459 | 135110,8900 | 356733,4400 | 240524,444205 | 69901,1677945 |
| Município Minerador | 188 | 0,1200 | 98322,0900 | 4157,214149 | 14734,2370858 |

\*Variáveis removidas do modelo pela baixa frequência apresentada.

Já no modelo 2 totalizaram 19 variáveis, na qual 6 delas foram excluídas por apresentar casos insuficientes, sendo elas “População dos 50 mais populosos”, “Recursos Hídricos”, “Penitenciárias”, “Esportes”, “Turismo” e “Mínimo per capita”.

Tabela 7 – Comportamento das variáveis independentes para o Modelo 2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variável** | **Frequência** | **Mínimo** | **Máximo** | **Média** | **Desvio Padrão** |
| VAF | 153 | 24537,8600 | 42546955,7100 | 1360718,592810 | 5772813,9100753 |
| Área Geográfica | 153 | 18934,4200 | 491010,0400 | 111730,054183 | 92173,4406311 |
| População | 153 | 24602,4500 | 2187774,7500 | 161880,452222 | 273759,6643254 |
| População dos 50 mais populosos\* | 3 | 2516747,0700 | 3096580,5800 | 2778568,366667 | 293972,4292627 |
| Educação | 118 | 315,9100 | 761118,0900 | 293090,911186 | 121313,7888878 |
| Produção de Alimentos | 153 | 28382,5100 | 220267,2600 | 84471,959281 | 34861,8871130 |
| Patrimônio Cultural | 151 | 79,5500 | 335252,5700 | 97054,408742 | 57471,3060566 |
| Meio Ambiente (A+B+C) | 125 | 788,6900 | 749673,9100 | 103062,670880 | 115965,9926427 |
| Programa Saúde da Família | 153 | 11928,5700 | 288593,9400 | 66320,209608 | 45925,8329280 |
| Saúde per capita | 153 | 12858,4400 | 142595,7000 | 53377,651438 | 23306,7498365 |
| Receita Própria | 153 | 6835,4200 | 637944,5500 | 96544,654118 | 84040,2913974 |
| Cota Mínima | 153 | 389044,6700 | 477162,1800 | 428818,664902 | 36600,3984666 |
| Município Minerador | 60 | 0,1800 | 11955,8500 | 713,059833 | 2383,6979605 |
| Recursos hídricos\* | 11 | 18288,5700 | 202687,2300 | 75496,900000 | 58618,7556740 |
| Penitenciárias\* | 43 | 383,6500 | 78220,9200 | 15826,828605 | 16171,3788490 |
| Esportes\* | 46 | 12,4300 | 164435,3100 | 37717,730652 | 41757,4777456 |
| Turismo\* | 26 | 201,4500 | 198902,0700 | 90227,736923 | 62281,1157005 |
| ICMS solidário | 153 | 40990,0200 | 3668295,2600 | 270054,491307 | 456888,3477342 |
| Mínimo per capita\* | 23 | 239,6600 | 338463,0100 | 179769,690435 | 104401,9773874 |

\*Variáveis removidas do modelo pela baixa frequência apresentada.

* 1. **CORRELAÇÃO ENTRE AS VARIÁVEIS INDEPENDENTES**

Analisar a correlação entre as variáveis independentes é de extrema importância antes de estimar um modelo de regressão, pois se as variáveis apresentarem alta correlação entre elas, as inferências baseadas no modelo podem apresentar erros e até mesmo serem de pouca confiabilidade. Assim, utilizando a correlação de Spearman, visto que as variáveis são ordinais, têm-se os seguintes coeficientes de correlação conforme mostra a Tabela 8.1, 8.2, 8.3 e 8.4.

* 1. **SELEÇÃO DAS VARIÁVEIS**

Para a estimação do modelo de regressão realizou-se o método *Backward Stepwise* para seleção das variáveis significantes de cada modelo, conforme mostram as tabelas.

Tabela 9.1 – Variável dependente PIB-M Tabela 9.2 – Variável dependente IFRS

|  |  |
| --- | --- |
| **PIB-M** | |
| **MODELO 1** | **MODELO 2** |
| VAF | VAF |
| População | Produção de Alimentos |
| Produção de Alimentos | Receita Própria |
| Município Minerador | Município Minerador |
| Receita Própria | ICMS solidário |

|  |  |
| --- | --- |
| **IFRS** | |
| **MODELO 1** | **MODELO 2** |
| Área Geográfica | Saúde per capita |
| População |  |
| Educação |  |
| Saúde per capita |  |
| Receita Própria |  |

Tabela 9.3 – Variável dependente IMRS Tabela 9.4 – Variável dependente IDH

|  |  |
| --- | --- |
| **IMRS** | |
| **MODELO 1** | **MODELO 2** |
| Área Geográfica | Área Geográfica |
| Produção de Alimentos | Patrimônio Cultural |
| Meio Ambiente | Receita Própria |
| Cota Mínima |  |

|  |
| --- |
| **IDH** |
| **MODELO 1** |
| VAF |
| Área Geográfica |
| População |
| Meio Ambiente |
| Programa Saúde da Família |

Feita a seleção das variáveis, utilizou-se o método VIF para excluir as variáveis com multicolinearidade, resultando assim na exclusão das seguintes variáveis:

* “População” do modelo 1 referente à variável dependente PIB;
* “ICMS Solidário” do modelo 2 referente à variável dependente PIB;
  1. **MODELOS DE REGRESSÃO**
     1. **MODELO PIB-M**

Tabela 10.1 – Estimativa dos parâmetros do Modelo 1 para a variável PIB-M.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | -1858,643 | 0,707 |  |
| VAF | -0,013 | < 0,001\* | 4,555 |
| Produção de Alimentos | 0,260 | < 0,001\* | 1,330 |
| Receita Própria | 1,192 | < 0,001\* | 4,817 |
| Município Minerador | 1,085 | < 0,001\* | 1,009 |

\* Valor significativo ao nível de 5% de significância.

O modelo 1 que refere-se à variável dependente PIB-M nos anos de 2002 a 2010 se mostrou adequado. Percebemos que a constante do modelo não apresentou significância, deste modo cabe ao pesquisador entender sua importância no modelo em questão. Como estudamos uma variável que não possuí, em sua forma primária, valores negativos, não há sentido em mantê-la. Já as variáveis independentes se mostraram significantes na construção da nossa variável dependente. Podemos concluir que a cada unidade do VAF, nosso PIB decresce em 0.013, o que pode ser resultado da Lei Robin Hood, que visa a distribuição do ICMS às cidades também consideradas pobres. As demais variáveis independentes que compõe o nosso modelo também se mostraram de peso na definição das outras dependentes, o que mostra que são importantes para o desenvolvimento do Vale do Jequitinhonha.

Tabela 10.2 – Estimativa dos parâmetros do Modelo 2 para a variável PIB-M.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | -60480,215 | 0,041\* |  |
| VAF | -0,019 | < 0,001\* | 4,177 |
| Produção de Alimentos | 0,647 | 0,009\* | 1,140 |
| Receita Própria | 1,578 | < 0,001\* | 4,050 |
| Município Minerador | 11,285 | 0,004\* | 1,078 |

\* Valor significativo ao nível de 5% de significância.

Após a reformulação do Lei Robi Hood, em 2011, algumas variáveis foram incluídas no cálculo para a distribuição do ICMS. Na construção deste segundo modelo percebemos que o PIB-M continua sendo explicado pelo mesmo grupo de variáveis, logo podemos concluir que esta modificação da Lei não alterou o comportamento do PIB na região estudada.

* + 1. **MODELO IRFS**

Tabela 11.1 – Estimativa dos parâmetros do Modelo 1 para a variável IRFS.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | 0,438540996 | < 0,001\* |  |
| Área Geográfica | 0,000000132 | < 0,001\* | 1,586687 |
| População | 0,000000022 | 0,2267702 | 2,668129 |
| Educação | 0,000000014 | 0,6198703 | 1,429266 |
| Produção de Alimentos | 0,000000059 | 0,4497052 | 1,759228 |
| Saúde per capita | 0,000000976 | < 0,001\* | 1,244379 |
| Receita Própria | -0,000000073 | 0,1614305 | 3,056025 |

\* Valor significativo ao nível de 5% de significância.

Estudando a variável IRFS como nossa dependente, percebemos que a mesma é formada basicamente pela constante. Podemos induzir que este fato se dá devido à linearidade do índice nas diferentes cidades do Vale, ou seja, apesar de sua variação todas seguem um padrão. Além da constante podemos concluir que o IRFS sofre um pequeno impacto da área geográfica e da saúde per capita. As demais variáveis não se mostraram significantes para o modelo.

Tabela 11.2 – Estimativa dos parâmetros do Modelo 2 para a variável IRFS.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | 0,486 | < 0,001\* |  |
| Saúde per capita | 3,770E-07 | 0,133 | 1,000 |

\* Valor significativo ao nível de 5% de significância.

Já após a reformulação da Lei em 2011, percebemos que a importância da constante em sua composição se acentuou, retirando as demais variáveis do modelo. A variável dependente “Saúde per capita” explica o restante da variável, apesar de não ter se mostrado significante.

* + 1. **MODELO IMRS**

Tabela 12.1 – Estimativa dos parâmetros do Modelo 1 para a variável IMRS.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | 0,552804144 | < 0,001\* |  |
| Área Geográfica | 0,000000205 | < 0,001\* | 1,466378677 |
| Produção de Alimentos | -0,000000447 | < 0,001\* | 1,522952114 |
| Meio Ambiente (A+B+C) | 0,000000148 | < 0,001\* | 1,05475786 |

\* Valor significativo ao nível de 5% de significância.

Não diferente do dito anteriormente, o IMRS das cidades pertencentes à região estudada seguem um “padrão”, ou seja, apesar de sua variação podemos dizer que em sua maioria seguem o mesmo nível. Este fato pode ser explicado devido à constante ter grande impacto no modelo, sofrendo pequenos ajustes pelas variáveis “área geográfica”, “produção de alimentos” e “meio ambiente”.

Tabela 12.2 – Estimativa dos parâmetros do Modelo 2 para a variável IMRS.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | 0,524506950 | < 0,001\* |  |
| Área Geográfica | -0,000000019 | 0,737603832 | 1,196744454 |
| Patrimônio Cultural | 0,000000276 | 0,010402759\* | 1,186409537 |
| Receita Própria | -0,000000063 | 0,345224112 | 1,123311412 |

\* Valor significativo ao nível de 5% de significância.

Após a reformulação da Lei, a constante permaneceu com grande impacto em sua definição, enquanto a variável “Área geográfica” perdeu sua significância, dando espaço à variável “Patrimônio Cultural”.

* + 1. **MODELO IDH-M**

Tabela 13.1 – Estimativa dos parâmetros do Modelo 2 para a variável IDH.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variáveis** | **Coef.** | **p-valor** | **VIF** |
| Constante | 0,592331713 | < 0,001\* |  |
| VAF | 0,000000015 | 0,036827117\* | 1,37831537 |
| Área Geográfica | -0,000000003 | 0,971866773 | 3,057620751 |
| População | 0,000000273 | 0,065296417 | 7,205144464 |
| Meio Ambiente | 1,57774E-07 | 0,005126255\* | 1,085091007 |
| Programa Saúde da Família | -4,17292E-07 | 0,145897023 | 5,016161587 |

\* Valor significativo ao nível de 5% de significância.

Por último, estudamos os impactos da Lei no IDH-M, variável esta que se apresentou com poucos casos, o que dificultou a construção do modelo. Nossa constante continua explicando grande parte da nossa dependente, sofrendo pequenos ajustes pela variável “VAF” e “Meio Ambiente”, sendo estas as únicas que se mostraram significativas.

* + 1. **Conclusão**

Apesar do sucesso na construção de todos os modelos, percebemos que nem todos se adequem à realidade. A variável dependente PIB, de fato, sofre uma grande influência das quatro variáveis descritas. É importante ser dito que duas, das quatro variáveis que compõe o modelo, eram utilizadas antes mesmo da existência da Lei Robin Hood. As variáveis dependentes IRFS, IMRS e IDH são compostas basicamente pela constante, sofrendo pequena influência de outras variáveis. Com isso podemos concluir que as cidades do vale, apesar de sua variação, se comportam de forma similar sem grande influência externa. A variável Produção de Alimentos” participou, de forma geral, na maioria dos modelos. Podemos concluir que a mesma se mostrou importante para o desenvolvimento do Vale.

* + 1. **Limitações**

Modelos de regressão linear são melhores definidos e mais confiáveis quando trabalhados sob um grande volume de casos, algo que não se apresentou no trabalho estudado. A falta de informações em algumas cidades comprometeu sua utilização na construção dos modelos.

**REFERÊNCIAS**

ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. Estatística aplicada à administração e economia. 2. Ed. São Paulo: Thomson Learning, 2007.

BARTLETT, M. S. Properties of sufficiency and statistical tests. Proceedings of the Royal Society of London, serie A, London, 1937.

BERK, K. N.Tolerance and condition in regression computations, Journal of the American Statistical Association, 72 (360), 863-866, 1977.

CONAB, Companhia Nacional de Abastecimento. Custos de produção - Culturas de Verão - Série Histórica. Disponível em: http://www.conab.gov.br/conteudos.php?a=1555&t=2. Acesso em: 6 de abril de 2016.

CORRAR, S. L.; PAULO, E.; DIAS FILHO, J.M. (Coord.). Análise multivariada para cursos de administração, ciências contábeis e economia. São Paulo: Atlas, 2007, 539p.

CHARNET, R. et al. Análise de modelos de regressão linear: com aplicações.2.ed. Campinas, São Paulo: Editora da Unicamp, 2008.

DRAPER, N.R.; SMITH, H. Applied regression analysis. 3.ed. New York: John Wiley e Sons, 1998. 706p.

FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L.; CHAN, B. L. Análise de dados:modelagem multivarida para tomada de decisões. Rio de Janeiro: Elsevier, 2009.

GUJARATI, D. N. Econometria básica. São Paulo: Makron Books, 2000.

HAIR, J. F.; ANDERSON, R.E.; TATHAM, R. L.; BLACK, W. C.; BABIN, B. J.Análise multivariada de dados. 5. Ed. Porto Alegre: Bookman, 2005.

HANSEN, D. R.; MOWEN, Maryanne M. Cost management. 2. Ed. Edition, Cincinnati, Ohio: South-Western CollegePublications, 1997.

HOFFMAN, R.; VIEIRA, S. Análise de regressão: uma introdução à econometria. 2. Ed. São Paulo: Câmara Brasileira do Livro, 1977.

MARTINS, E.; Contabilidade de custos: Inclui o ABC. 8. Ed. São Paulo: Atlas, 2006.

MATOS. O.C. (2000). Econometria Básica. São Paulo: Atlas, p. 124.

MORETTIN. P. A.; TOLOI, C. M. C. Séries temporais. 2. Ed., Atual: São Paulo, 1987.

SANTOS, J. J. Contabilidade e Análise de custos: Modelo contábil, Método de depreciação, ABC – Custeio Baseado em Atividades, análise atualizada de encargos sociais sobre salários. 5. Ed. São Paulo: Atlas, 2009.

SANTOS, J. L.; SCHMIDT, P., & PINHEIRO, P. R. Fundamentos de Gestão Estratégica de Custos. São Paulo: Atlas S/A. 2006.

SEBER, G. A. F. Linear Regression Analysis. New York: Wiley, 1997.

SHAPIRO, S. S.; WILK, M. B. An Analysis of Variance Test for Normality. Biometrika, Vol. 52, pp. 591-611, 1965.

KUTNER, M. H. et al. Applied linear models. 5th ed. New York: McGraw-Hill Irwin, 2004.

TAMHANE, A.; DUNLOP.D. Statistics and data analysis: From Elementary to Intermediate. Prentice Hall, 2000.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Tabela 8.1 – Correlação de Spearman para as variáveis estudadas no Modelo 1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| VAF | 1,000 |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Área Geográfica | 0,687 | 1,000 |  |  |  |  |  |  |  |  |  |  |  |  |
| População | 0,749 | 0,791 | 1,000 |  |  |  |  |  |  |  |  |  |  |  |
| População dos 50 mais populosos | 1 | 1,000\*\* | 1 | 1,000 |  |  |  |  |  |  |  |  |  |  |
| Educação | 0,297 | 0,516 | 0,503 | 0,310 | 1,000 |  |  |  |  |  |  |  |  |  |
| Produção de Alimentos | 0,471 | 0,645 | 0,753 | 1 | 0,574 | 1,000 |  |  |  |  |  |  |  |  |
| Mateus Leme/Mesquita | 0,266 | 0,260 | 0,088 |  | 0,077 | -0,149 | 1,000 |  |  |  |  |  |  |  |
| Patrimônio Cultural | 0,265 | 0,25 | 0,202 | -0,367 | 0,027 | 0,109 | 0,699 | 1,000 |  |  |  |  |  |  |
| Meio Ambiente (A+B+C) | -0,241 | -0,236 | -0,223 |  | -0,187 | -0,154 | -0,400 | 0,057 | 1,000 |  |  |  |  |  |
| Programa Saúde da Família | 0,589 | 0,671 | 0,77 | 0,817 | 0,537 | 0,597 | -0,053 | 0,195 | -0,106 | 1,000 |  |  |  |  |
| Saúde per capita | -0,185 | -0,174 | -0,263 | 0,467 | 0,114 | -0,102 | 0,164 | 0,090 | 0,419 | -0,014 | 1,000 |  |  |  |
| Receita Própria | 0,638 | 0,604 | 0,685 | 0,917 | 0,422 | 0,563 | 0,199 | 0,22 | -0,045 | 0,662 | 0,179 | 1,000 |  |  |
| Cota Mínima | 0,264 | 0,381 | 0,373 | 1 | 0,639 | 0,495 | 0,490\* | 0,102 | 0,187 | 0,518 | 0,567 | 0,627 | 1,000 |  |
| Município Minerador | 0,272 | 0,192 | 0,105 | 1 | -0,114 | -0,042 | 1 | 0,402 | 0,189 | 0,128 | 0,249 | 0,268 | 0,118 | 1,000 |

Tabela 8.2 – Tabela de -valores da correlação de Spearman para as variáveis estudadas no Modelo 1.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| VAF |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Área Geográfica | < 0,001\* |  |  |  |  |  |  |  |  |  |  |  |  |
| População | < 0,001\* | < 0,001\* |  |  |  |  |  |  |  |  |  |  |  |
| População dos 50 mais populosos |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Educação | < 0,001\* | < 0,001\* | < 0,001\* | 0,456 |  |  |  |  |  |  |  |  |  |
| Produção de Alimentos | < 0,001\* | < 0,001\* | < 0,001\* |  | < 0,001\* |  |  |  |  |  |  |  |  |
| Mateus Leme/Mesquita | 0,243 | 0,256 | 0,703 |  | 0,760 | 0,518 |  |  |  |  |  |  |  |
| Patrimônio Cultural | < 0,001\* | < 0,001\* | < 0,001\* | 0,332 | 0,589 | 0,020\* | < 0,001\* |  |  |  |  |  |  |
| Meio Ambiente (A+B+C) | < 0,001\* | < 0,001\* | 0,002\* |  | 0,021\* | 0,034\* | 0,600 | 0,440 |  |  |  |  |  |
| Programa Saúde da Família | < 0,001\* | < 0,001\* | < 0,001\* | 0,007\* | < 0,001\* | < 0,001\* | 0,836 | < 0,001\* | 0,148 |  |  |  |  |
| Saúde per capita | < 0,001\* | < 0,001\* | < 0,001\* | 0,205 | 0,024\* | 0,029\* | 0,478 | 0,056 | < 0,001\* | 0,772 |  |  |  |
| Receita Própria | < 0,001\* | < 0,001\* | < 0,001\* | < 0,001\* | < 0,001\* | < 0,001\* | 0,388 | < 0,001\* | 0,534 | < 0,001\* | < 0,001\* |  |  |
| Cota Mínima | < 0,001\* | < 0,001\* | < 0,001\* |  | < 0,001\* | < 0,001\* | 0,024\* | 0,029\* | 0,010\* | < 0,001\* | < 0,001\* | < 0,001\* |  |
| Município Minerador | < 0,001\* | 0,008\* | 0,151 |  | 0,134 | 0,564 |  | < 0,001\* | 0,101 | 0,087 | < 0,001\* | < 0,001\* | 0,107 |