



**Texto n. 31**

**Textos para Discussão**  
**ISSN – 2447-8210**

**Modelagem matemática da  
frota de automóveis de  
Minas Gerais por regressão  
linear e uso de *softwares*  
livres**

**Lucas Morais Melo**

## MODELAGEM MATEMÁTICA DA FROTA DE AUTOMÓVEIS DE MINAS GERAIS POR REGRESSÃO LINEAR E USO DE *SOFTWARES* LIVRES

Lucas Morais Melo<sup>1</sup>

### RESUMO

Este trabalho apresenta a criação de um modelo matemático que seja capaz de realizar previsões a respeito do quantitativo de automóveis no Estado de Minas Gerais, baseando-se nos dados do Departamento Nacional de Trânsito (DENATRAN). Para isso, sendo utilizada a técnica de regressão linear e contando com o auxílio de *softwares* livres e gratuitos. A diferença percentual média entre os valores oficiais e os valores obtidos pelo modelo, no período de 2009 a 2019, foi de apenas 0,0006%. Para validação do modelo foi utilizado o valor da frota de automóveis para o mês de janeiro de 2020, valor fornecido pelo DENATRAN, resultando-se em uma diferença percentual de 0,0079% para o valor estimado pelo modelo, mostrando-se, assim, que o modelo se ajusta muito bem em relação aos dados oficiais.

**Palavras-chave:** Modelagem Matemática. Regressão linear. Software livre.

### ***MATHEMATICAL MODELING OF THE MINAS GERAIS AUTOMOBILE FLEET BY LINEAR REGRESSION AND USE OF FREE SOFTWARE***

### **ABSTRACT**

This work presents the creation of a mathematical model that is capable of making predictions regarding the quantity of automobiles in the State of Minas Gerais, based on data from the National Traffic Department (DENATRAN). For this, using the linear regression technique and relying on free and free software. The average percentage difference between the official values and the values obtained by the model, in the period from 2009 to 2019, was only 0.0006%. To validate the model, the value of the car fleet for the month of January 2020 was used, a value provided by DENATRAN, resulting in a percentage difference of 0.0079% to the value estimated by the model, thus showing that the model fits very well in relation to official data.

**Keywords:** *Mathematical modeling. Linear regression. Free software.*

---

<sup>1</sup> Bacharel em Engenharia Elétrica pela Universidade Federal de São João del-Rei (UFSJ). Graduando em Licenciatura em Matemática pelo Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais – Campus Formiga (IFMG). lucasmmelolp@gmail.com

## 1 INTRODUÇÃO

Localizado na região sudeste do Brasil, Minas Gerais, com uma população estimada em aproximadamente 21 milhões de habitantes, é o segundo estado mais populoso do país, atrás apenas do Estado de São Paulo. Com 853 municípios, Minas Gerais é o estado com o maior número de municípios do país. Em relação à área da unidade territorial, com 586.521,121 Km<sup>2</sup>, Minas Gerais é o quarto estado com maior área territorial. Com um Produto Interno Bruto (PIB) de R\$ 287 bilhões, Minas Gerais é um dos estados mais importantes para a economia brasileira (ESTADO DE MINAS GERAIS, 2020; IBGE, 2020).

Os dados anteriormente citados mostram a importância, posição de destaque no cenário nacional e algumas peculiaridades geográficas e demográficas do Estado de Minas Gerais. A utilização de automóveis em um estado com essas características torna-se natural.

Visando estimar a frota de automóveis de Minas Gerais, é desenvolvido neste artigo um modelo matemático que seja capaz de realizar estimativas desta frota para os próximos anos.

Para o desenvolvimento deste modelo foram utilizados dados oficiais do Departamento Nacional de Trânsito (DENATRAN). Além disso, foi utilizada a técnica de regressão linear, com a finalidade de obter uma curva que se ajuste bem aos dados reais. Para auxílio gráfico, foi utilizado o *software* LibreOffice Calc e a linguagem de programação GNU Octave, ambos recursos computacionais livres e gratuitos.

## 2 DESENVOLVIMENTO

### 2.1 Regressão linear

A técnica de regressão “consiste na realização de uma análise estatística com o objetivo de verificar a existência de uma relação funcional entre uma variável dependente com uma ou mais variáveis independentes” (PETERNELLI, 2005, p. 3).

A regressão linear é uma técnica de Modelagem Matemática que comumente baseia-se no Método dos Mínimos Quadrados (MMQ). Esta técnica permite que seja obtida a curva ótima de ajuste em relação aos pontos observados, em outras

palavras, por meio da técnica de regressão linear é possível determinar qual é a melhor curva que se ajusta em relação aos pontos de um diagrama de dispersão, isto quando outros tipos de ajustes não são melhores, ajustes como, por exemplo, o ajuste logístico ou mesmo ajustes de regressão polinomial que tenham grau maior do que 1, que eventualmente podem se ajustar melhor a depender do diagrama de dispersão.

Basicamente, o MMQ consiste em determinar parâmetros de uma função, de modo que seja minimizada a soma dos quadrados dos desvios entre os valores observados (reais) e os valores ajustados (obtidos pelo modelo).

O ajuste linear é da forma de uma função afim, ou seja,  $f(x) = ax + b$ ;  $a, b \in \mathbb{R}$ , onde  $\mathbb{R}$  é o conjunto dos números reais. Sendo  $(\bar{x}_i, \bar{y}_i)$ ,  $i = 1, 2, 3, \dots, n$  os pontos observados (reais), com o auxílio do MMQ, e também de técnicas de Cálculo Diferencial em algumas passagens, chega-se às seguintes fórmulas para determinação dos coeficientes  $a$  (Equação 1) e  $b$  (Equação 2):

$$a = \frac{\sum \bar{x}_i \bar{y}_i - n \bar{x} \bar{y}}{\sum \bar{x}_i^2 - n \bar{x}^2} \quad (1)$$

$$b = \bar{y} - a \bar{x} \quad (2)$$

onde:  $\bar{y} = \frac{\sum \bar{y}_i}{n}$  e  $\bar{x} = \frac{\sum \bar{x}_i}{n}$ .

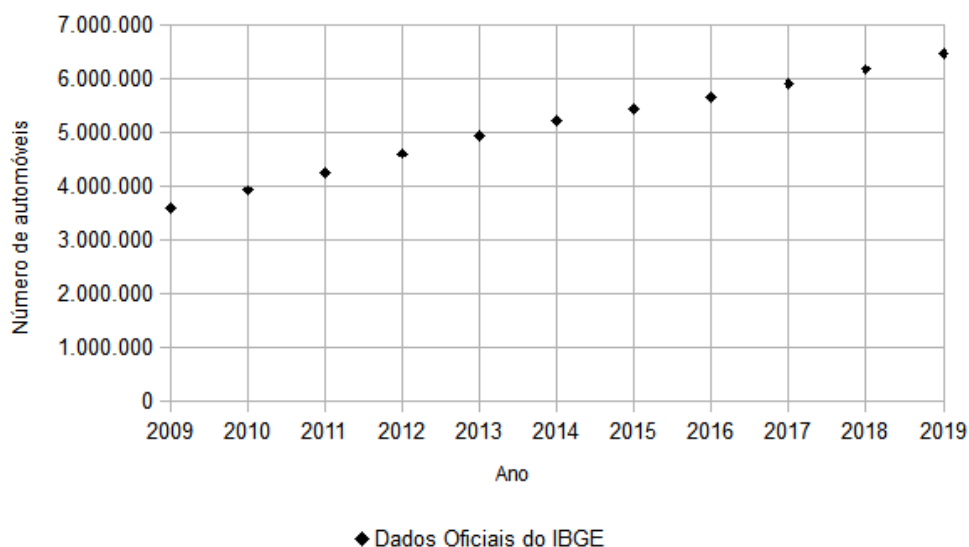
## 2.2 Frota de automóveis em Minas Gerais

Na Tabela 1 é possível ver a quantidade de automóveis em Minas Gerais no período entre 2009 e 2019.

Ano	Número de automóveis
2009	3.600.042
2010	3.922.908
2011	4.252.225
2012	4.602.143
2013	4.926.454
2014	5.215.337
2015	5.441.609
2016	5.652.316
2017	5.901.355
2018	6.169.791
2019	6.467.501

**Tabela 1.** Frota de automóveis em Minas Gerais de 2009 a 2019  
Fonte: DENATRAN (2020a).

Com o auxílio do *software* LibreOffice Calc, pode-se plotar os dados da Tabela 1 e obter o diagrama de dispersão apresentado no Gráfico 1.



**Gráfico 1.** Diagrama de dispersão da frota de automóveis de Minas Gerais de 2009 a 2019  
Fonte: Elaborado pelo autor.

### 2.3 Modelo matemático e resultados

Neste caso particular, tem-se que:

$$\bar{x}_1 = 2009, \bar{x}_2 = 2010, \bar{x}_3 = 2011, \bar{x}_4 = 2012, \bar{x}_5 = 2013, \bar{x}_6 = 2014, \bar{x}_7 = 2015, \\ \bar{x}_8 = 2016, \bar{x}_9 = 2017, \bar{x}_{10} = 2018, \bar{x}_{11} = 2019$$

e que:  $\bar{y}_1 = 3.600.042$ ,  $\bar{y}_2 = 3.922.908$ ,  $\bar{y}_3 = 4.252.225$ ,  $\bar{y}_4 = 4.602.143$ ,  
 $\bar{y}_5 = 4.926.454$ ,  $\bar{y}_6 = 5.215.337$ ,  $\bar{y}_7 = 5.441.609$ ,  $\bar{y}_8 = 5.652.316$ ,  $\bar{y}_9 = 5.901.355$ ,  
 $\bar{y}_{10} = 6.169.791$ ,  $\bar{y}_{11} = 6.467.501$ . Deste modo, segue que:  $\bar{x} = 2014$  e  
 $\bar{y} = 5.104.698,2727$ . Como de 2009 a 2019 tem-se 11 termos, segue que  $n = 11$ .

Portanto, pode-se determinar o valor do coeficiente  $a$  pela Equação (1):

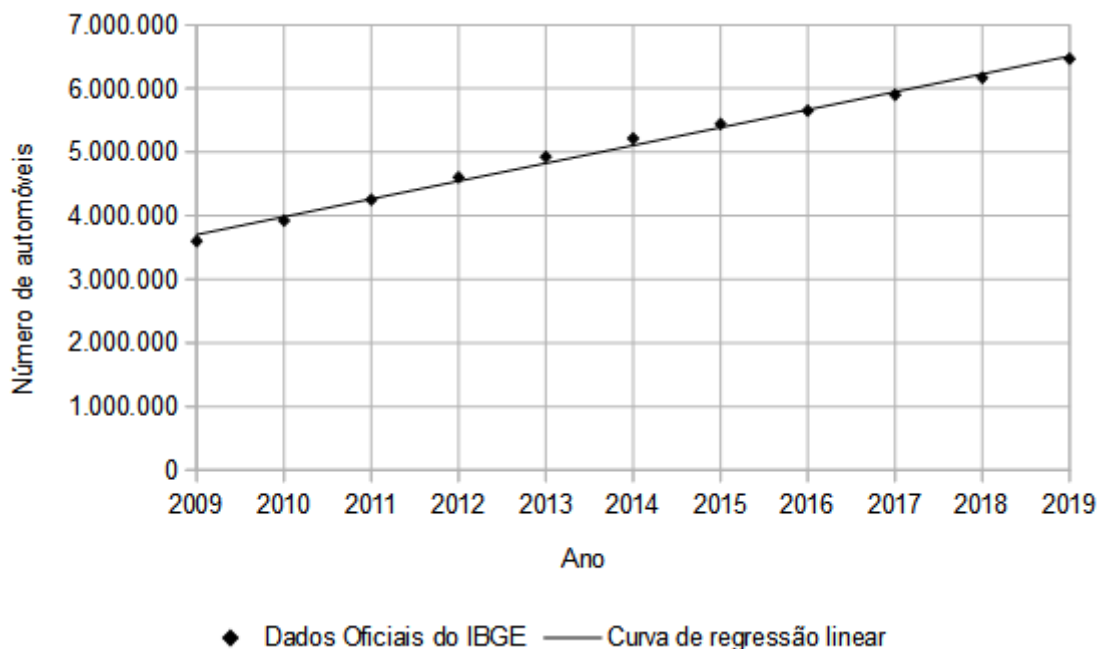
$$a = \frac{\sum \bar{x}_i \bar{y}_i - n \bar{x} \bar{y}}{\sum \bar{x}_i^2 - n \bar{x}^2} = \frac{113.120.373.252 - 11 \cdot 10.280.862.321,273}{44.618.266 - 11 \cdot 4.056.196} \cong 280.797,4364$$

Após ser determinado o coeficiente  $a$ , pela Equação (2), determina-se o valor do coeficiente  $b$ :

$$b = \bar{y} - a \bar{x} = 5.104.698,2727 - 280.797,4364 \cdot 2014 \cong -560.421.338,5639$$

logo, a reta de regressão é:  $f(x) = 280.797,4364x - 560.421.338,5639$ .

O Gráfico 2 apresenta a reta de regressão  $f$  ajustada em relação ao gráfico de dispersão. O Gráfico 2 foi obtido por meio de ferramenta específica do LibreOffice Calc.



**Gráfico 2.** Diagrama de dispersão e reta de regressão da frota de automóveis de Minas Gerais de 2009 a 2019  
Fonte: Elaborado pelo autor.

De acordo com Peternelli (2005), o coeficiente de correlação de Pearson ( $r$ ) é utilizado para mensurar o grau de associação entre duas variáveis. “Se for observada uma associação entre as variáveis quantitativas (a partir de um diagrama de dispersão, por exemplo), é muito útil quantificar essa associabilidade” (PETERNELLI, 2005, p. 1). Segundo Subramanian, Coutinho e Silva (2007), para valores de  $r$  próximos de 1 ou de -1, considera-se que a correlação é forte, ou seja, quanto mais próximo de 0 mais fraca é a correlação. Deste modo, é interessante calcular o valor de  $r$  com a finalidade de obter um indicador do quão bem a curva ajusta-se ao gráfico de dispersão.

Segundo Galarça *et al* (2010), calcula-se  $r$  pela fórmula:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Deste modo, para o caso particular deste trabalho:

$$r = \frac{30.887.718}{\sqrt{(110) \cdot (8.723.612.924.998,18)}} \cong 0,99422$$

Mostrando-se, assim, que existe uma forte correlação entre o gráfico de dispersão e a reta de regressão.

Para o cálculo dos valores previstos pelo modelo matemático obtido, utilizou-se a linguagem de programação GNU Octave.

Na Tabela 2 é possível visualizar a quantidade de carros prevista pelo modelo para cada ano, durante o período de 2009 a 2019. Além disso, a Tabela 2 apresenta a diferença percentual entre os valores oficiais dados pelo DENATRAN (Nº auto. DENATRAN) e os valores obtidos por meio do modelo (Nº auto. modelo). A média aritmética das diferenças percentuais apresentadas na Tabela 2 é de apenas 0,0006%.

Ano	Nº auto. DENATRAN	Nº auto. modelo	Diferença
2009	3.600.042	3.700.711	0,0014%
2010	3.922.908	3.981.509	0,0007%
2011	4.252.225	4.262.306	0,0001%
2012	4.602.143	4.543.103	0,0006%
2013	4.926.454	4.823.901	0,0010%
2014	5.215.337	5.104.698	0,0011%
2015	5.441.609	5.385.496	0,0005%
2016	5.652.316	5.666.293	0,0001%
2017	5.901.355	5.947.091	0,0004%
2018	6.169.791	6.227.888	0,0005%
2019	6.467.501	6.508.685	0,0003%

**Tabela 2.** Número de automóveis em Minas Gerais entre 2009 e 2019 obtidos pelo modelo e diferenças percentuais  
Fonte: Elaborado pelo autor.

Para validação do modelo matemático foi utilizada a tabela que fornece a quantidade da frota de automóveis em Minas Gerais para o mês de janeiro de 2020, elaborada por (DENATRAN, 2020b). De acordo com este documento no mês de janeiro de 2020, o Estado de Minas Gerais possuía 6.481.149 automóveis. Para o cálculo da frota deste mesmo período por meio do modelo utilizou-se  $x = 2019,0833$ , ou seja, foi calculado o valor de  $f(2019,0833)$ . Deste modo, de acordo com o modelo desenvolvido por este trabalho, o total da frota de automóveis em Minas Gerais no mês de janeiro de 2020 é de 6.532.084. Assim sendo, chega-se a diferença percentual, entre o valor oficial do DENATRAN e o valor obtido pelo modelo, de apenas 0,0079%.

Portanto, por meio do modelo matemático desenvolvido neste trabalho, estima-se que a quantidade de automóveis no Estado de Minas Gerais no final do ano de 2020 será de 6.789.482; já para o ano de 2025, ter-se-iam 8.193.470 automóveis.

### 3 CONSIDERAÇÕES FINAIS

Verificou-se que o modelo matemático obtido ajusta-se muito bem em relação aos dados oficiais do DENATRAN, para isso, utilizou-se o valor do coeficiente de correlação de Pearson e o erro percentual médio. Para a validação do modelo foi utilizado um dado mensal, também elaborado pelo DENATRAN, chegando-se a uma pequena diferença percentual entre o dado oficial e o estimado pelo modelo. Deste modo, acredita-se que o modelo desenvolvido neste trabalho possa realizar estimativas consideráveis em períodos próximos.

Obviamente o modelo descrito neste trabalho torna-se não recomendável para a estimação de valores em longo prazo, visto que neste caso o modelo poderia retornar valores que tenderiam ao infinito. Deste modo, sugere-se que para efetiva obtenção de previsões em longo prazo seja desenvolvido um modelo que envolva outras variáveis, tais como demográficas e econômicas.

Espera-se que este trabalho sirva como um exemplo prático para o ensino de técnicas de ajuste de curvas, pois, como visto aqui, por meio de uma técnica matemática simples é possível chegar a resultados interessantes, contextualizando o processo de ensino e aprendizagem por meio da abordagem de um problema do mundo real.

### REFERÊNCIAS

ESTADO DE MINAS GERAIS. **Dados Gerais**. Disponível em: <<https://www.mg.gov.br/conteudo/conheca-minas/geografia/dados-gerais>>. Acesso em: 2 mar. 2020.

DEPARTAMENTO NACIONAL DE TRÂNSITO – DENATRAN. **Estatísticas Frota de Veículos - DENATRAN**. Disponível em: <<https://www.infraestrutura.gov.br/component/content/article/115-portal-denatran/8552-frota-de-veiculos-2019.html>>. Acesso em: 2 mar. 2020a.



\_\_\_\_\_. **Frota de Veículos – 2020.** Disponível em: <  
<https://infraestrutura.gov.br/component/content/article/115-portal-denatran/9484-frota-de-veiculos-2020.html>>. Acesso em: 1 abr. 2020b.

GALARÇA, S. P. et al. Correlação de Pearson e análise de trilha identificando variáveis para caracterizar porta-enxerto de *Pyrus communis* L. **Ciênc. agrotec.**, Lavras, v. 34, n. 4, p. 860-869, jul./ago. 2010. Disponível em: <  
<http://www.scielo.br/pdf/cagro/v34n4/v34n4a10.pdf>>. Acesso em: 3 mar. 2020.

INSTITUTO BRASILEIRO DE ESTATÍSTICA E GEOGRAFIA – IBGE. **Cidades e Estados.** Disponível em: <  
<https://www.ibge.gov.br/cidades-e-estados/mg.html>>. Acesso em: 2 mar. 2020.

PETERNELLI, L. A. **Estatística I.** Apostila para cursos da Universidade Federal de Viçosa, 2005. Disponível em: <  
[www.dpi.ufv.br/~peterneli/inf162.www.16032004/materiais/CAPITULO9.pdf](http://www.dpi.ufv.br/~peterneli/inf162.www.16032004/materiais/CAPITULO9.pdf)>. Acesso em: 10 fev. 2020.

SUBRAMANIAN, A.; COUTINHO, A. S.; SILVA, L. B. Aplicação de método e técnica multivariados para previsão de variáveis termoambientais e perceptivas. **Produção**, v. 17, n. 1, p. 052-070, jan./abr. 2007. Disponível em: <  
<http://www.scielo.br/pdf/prod/v17n1/03.pdf>>. Acesso em: 3 mar. 2020.