



Universidade Federal de Uberlândia  
Faculdade de Matemática

Bacharelado em Estatística

**AVALIAÇÃO DO INDICADOR DE  
LEITÕES ENTREGUES VIA ANÁLISE  
DE REGRESSÃO**

Leonardo Pereira Borges

Uberlândia-MG

2019



**Leonardo Pereira Borges**

**AVALIAÇÃO DO INDICADOR DE  
LEITÕES ENTREGUES VIA ANÁLISE  
DE REGRESSÃO**

Trabalho de conclusão de curso apresentado à Co-  
ordenação do Curso de Bacharelado em Estatística  
como requisito parcial para obtenção do grau de  
Bacharel em Estatística.

Orientador: Professor Dr. Lúcio Borges de Araújo

**Uberlândia-MG**

**2019**





**Universidade Federal de Uberlândia  
Faculdade de Matemática**

**Coordenação do Curso de Bacharelado em Estatística**

A banca examinadora, conforme abaixo assinado, certifica a adequação deste trabalho de conclusão de curso para obtenção do grau de Bacharel em Estatística.

Uberlândia, \_\_\_\_\_ de \_\_\_\_\_ de 20\_\_\_\_\_

**BANCA EXAMINADORA**

---

Professor Dr. Lúcio Borges de Araújo

---

Professor Dr. Janser Moura Pereira

---

Professor Dr. José Waldemar da Silva

**Uberlândia-MG  
2019**



# AGRADECIMENTOS

A minha esposa e minhas filhas que junto comigo cursaram uma etapa da vida com bastante paciência, incentivo, dedicação e me deram o apoio necessário para que eu não desistisse e chegar onde hoje cheguei.

Aos meus pais que sempre estiveram do meu lado me incentivando e me ajudando durante essa etapa da minha vida, principalmente por acreditar no meu potencial.

A Universidade Federal de Uberlândia pelo discernimento de seus profissionais compartilhando seus conhecimentos e dando a base necessária para a realização deste trabalho.

Ao professor Dr. Lúcio Borges de Araújo pela orientação deste trabalho, pelo apoio, paciência e ensinamentos durante o desenvolvimento deste trabalho e as aulas ministradas durante as aulas de graduação.

Agradeço também aos meus amigos, colegas de trabalho e de faculdade pelo apoio e que sempre torceram por mim nesta jornada de graduação.



# RESUMO

Este trabalho tem como objetivo avaliar por meio da regressão linear múltipla os dados que influenciam no resultado da quantidade de leitões entregues por fêmea ao ano (LEFA) de uma empresa produtora de suínos. A pesquisa foi realizada com os dados que foram coletados durante o período de Jan/2018 a Dez/2018. São 142 amostras com 14 variáveis que compõem o estudo, cada amostra representa um produtor com uma quantidade média de 2.600 fêmeas, durante as análises foram selecionadas por meios de modelos estatísticos 4 variáveis para compor o modelo final de nosso estudo. Os produtores acreditam que o maior impacto no LEFA é a mortalidade do leitão durante o seu período de creche. Após os ajustes dos modelos concluiu-se que 4 variáveis influenciam no resultado sendo elas a mortalidade no desmame, mortalidade na creche, retorno ao cio e a taxa de parto. A variável que os produtores acreditam ser a de maior impacto, foi analisada e dentre as 4 variáveis final do modelo, esta variável foi a de menor impacto. Diante das análises realizadas ao longo deste trabalho, observa-se que há grandes oportunidades de tratativas que geram um impacto mais expressivo.

**Palavras-chave:** Regressão Linear Múltipla, LEFA, VIF, Stepwise .



# ABSTRACT

The job was average in regression the linear lectures of digital to an example (LEFA) of a pig producer. The research was conducted based on data collected from January 2018 to December 2018. There are 142 samples with 14 variables that make up the study, each series represents a print with an average quantity of 2.600 females, during the analyzes. were selected by the forms of 4-variable statistical models to compose the final model of our study. Producers believe that the biggest impact on LEFA is the piglet's mortality during its nursery period. The data of the completed completed 4 weeks influence in return, which are the mortality in weaning in a nursery. One variable that can be used as the highest impact was analyzed in relation to 4 final variables of the model, this variable was of lower impact. Faced with overworked writing, it is observed that there are great treatment opportunities that generate a more significant impact.

**Keywords:** Multiple Linear Regression, LEFA, VIF, Stepwise.



# SUMÁRIO

Lista de Figuras	I
Lista de Tabelas	III
<b>1</b> Introdução	<b>1</b>
<b>2</b> Materiais e Métodos	<b>3</b>
2.1 Dados . . . . .	3
2.2 Análise Estatística: Análise de Regressão . . . . .	3
2.3 Estimativa com o Método Mínimos Quadrados . . . . .	4
2.3.1 Representação Matricial do MRLM . . . . .	5
2.4 Multicolinearidade . . . . .	6
2.5 Método Stepwise . . . . .	7
<b>3</b> Resultados	<b>9</b>
<b>4</b> Conclusões	<b>19</b>
Referências Bibliográficas	<b>21</b>



# LISTA DE FIGURAS

3.1	Correlação entre variáveis consideradas no estudo . . . . .	10
3.2	Homoscedasticidade dos Resíduos . . . . .	12
3.3	Normalidade dos Resíduos . . . . .	13
3.4	Distribuição dos Resíduos . . . . .	14
3.5	Pontos Influentes no modelo . . . . .	15



# LISTA DE TABELAS

3.1	Estatísticas descritivas das variáveis a serem utilizadas no estudo . . . . .	9
3.2	Valores do VIF das variáveis explicativas selecionadas . . . . .	11
3.3	Análise de regressão das variáveis explicativas selecionadas . . . . .	11
3.4	Impactos financeiros, mensais e anuais, de acordo com a evolução do indicador LEFA . . . . .	16



# 1. INTRODUÇÃO

A suinocultura brasileira posicionada entre as cadeias produtivas mais avançadas do mundo, é uma gigantesca cadeia produtiva que gera renda e desenvolvimento pelo interior do país [1].

Segundo a [1], o Brasil é o 4º maior produtor de carne suína no mundo. E para uma empresa produtora de suínos ter o seu negócio rentável, precisa cuidar de todos os detalhes para que tenha animais com custo competitivo. Uma cadeia de produção de suínos em que o objetivo é a criação dos animais para abate, tem um longo ciclo e possui períodos em que o suíno precisa apresentar bons desempenhos para que seu desenvolvimento seja viável no processo.

No processo analisado, o suíno possui duas fases do nascimento ao abate. Primeira fase é representada pelos primeiros 60 dias de vida, essa fase é chamada de Sistema Produtor de Leitões (SPL), a segunda fase é o Sistema Vertical de Terminados (SVT), é quando os leitões são alojados até o abate, o segundo ciclo da fase é feito no período de 120 dias [5].

Para mensurar a eficiência do processo na fase do SPL temos o indicador de leões entregues no ano (LEFA) que é fundamental para verificar o seu desempenho. O LEFA é um indicador que possui uma meta mensal de 26,8 leitões entregues, porém têm vários fatores que influenciam em seu resultado, portanto sendo necessário todo o cuidado com os animais desde o seu nascimento [12].

Alguns fatores do processo pode influenciar no indicador LEFA, sendo eles: os nascidos totais que é a quantidade de leitões que nasceram sendo vivos ou mortos, nascidos vivos são aqueles leitões que nasceram vivos no parto, têm os natimortos que são os que nasceram mortos, mumificados são os que morreram durante a gestação e ficaram na fêmea até o momento do parto, mortos no desmame são os que morreram nos primeiros dias de vida, parto fêmea ano é a quantidade de partos que a fêmea deve fazer no ano, idade desmame é a idade em que o leitão foi separado da mãe, idade da creche é a idade que o animal foi tirado do processo do SPL e transferido para o SVT, peso da creche é o peso em que foi adquirido do nascimento até a transferência para o SVT, mortalidade creche é a quantidade de animais mortos desde o seu desmame até a sua transferência para o SVT, retorno ao cio é o tempo em que a fêmea após o parto teve o novo cio e foi inseminada, taxa de parto refere-se a quantidade de partos que a fêmea teve durante o ano, mortalidade da fêmea é quando a fêmea morre durante o seu período de vida como matriz e o plantel que é a quantidade de fêmeas no seu estado de reprodução. Os fatores acima são observados durante o seu processo de criação no SPL [5].

O SPL é um ciclo do processo muito importante no desenvolvimento do animal, este período em que fica alojado no SPL o leitão deve apresentar bom desempenho, pois dali em diante ele

vai levar seu bom ou ruim resultado para todo o seu ciclo de vida [12]. No ano de 2018 o sistema foi muito impactado pelo alto índice de mortalidade no seu início de vida e isso pode prejudicar todo o processo e o aumento de custo no animal.

In Loco foi observado e afirmado pelos produtores que a principal causa que está afetando o desempenho são as altas taxas de mortalidade na maternidade onde possui *diversos fatores motivadores*, porem a mortalidade na maternidade será transformado em indicadores no fim da fase do sistema, ou seja, nos seus 60 dias de idade. Observando este fato da alta mortalidade a atenção e trabalhos foram voltados para a identificação das variáveis que estão impactando o seu desempenho.

Na gestação da porca nascem em torno de 14 leitões, e no nascimento já se encontra alguns fatores que podem influenciar neste desempenho que é o peso de nascimento do leitão e a quantidade de colostro mamado [10].

Outro fator que impacta na mortalidade é a diarreia na fase inicial do leitão, durante os 21 dias de desmame, e as causas de diarreia pode ser por falta de leite na primeira semana, manejo de rotação de mamadas, presença de infecções bacterianas e outros protozoários como exemplo a coccidiose que pode apresentar no período de 72 horas pós nascimento [10].

O objetivo deste trabalho é verificar quais são as principais variáveis que interferem (melhoram ou pioram) nos resultados de entrega de leitões LEFA.

## 2. MATERIAIS E MÉTODOS

### 2.1 DADOS

Os dados utilizados neste trabalho foram coletados durante o período de janeiro/2018 a dezembro/2018 da região de Uberlândia-MG. São 142 amostras coletadas com 14 variáveis, cada amostra representa um produtor com um plantel médio de 2.600 fêmeas. O banco de dados para análise, é composto pela variável LEFA (que é a variável resposta) e as demais variáveis são as independentes, conforme mencionado abaixo:

$Y_i$ : Unidades Leitões entregues no ano por fêmea (LEFA);

$X_1$ : % Nascidos totais (NT);

$X_2$ : % Nascidos vivos (NV);

$X_3$ : % Natimortos (NM);

$X_4$ : % Mumificados (MM);

$X_5$ : % Mortos no Desmame (M Desm);

$X_6$ : % Parto fêmea ano (PFA);

$X_7$ : Média Idade de saída desmame (ID Desm);

$X_8$ : Média Idade de saída da creche (ID Creche);

$X_9$ : Média Peso (kg) de saída da creche (P Creche);

$X_{10}$ : % Mortalidade da creche (M Creche);

$X_{11}$ : % Retorno ao cio (R Cio);

$X_{12}$ : % Taxa de parto (TX Parto);

$X_{13}$ : % Mortalidade fêmea (M Fêmea);

$X_{14}$ : Quantidade do Plantel (Plantel);

Para a identificação das principais causas que impactam diretamente no processo, será aplicada algumas ferramentas e métodos estatísticos.

### 2.2 ANÁLISE ESTATÍSTICA: ANÁLISE DE REGRESSÃO

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 [6].

Tem-se uma regressão linear múltipla quando admitimos que o valor da variável dependente é função linear de duas ou mais variáveis independentes [7].

A análise de regressão múltipla tem por objetivo 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. 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 ([4] e [6]).

O modelo de regressão múltipla é dado por [3]:

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + u_i$$

para que:  $Y_i$  é o fenômeno em estudo (variável dependente);  $\beta_0$  representa o intercepto (constante);  $\beta_1, \beta_2, \dots, \beta_k$  são os coeficientes associados a cada variável independente (coeficientes angulares);  $(x_{ki})$  são as variáveis explicativas (independentes) com  $i=1, 2, \dots, n$ ;  $u_i$  é o termo do erro. O erro  $u_i$ , também chamado de resíduo, representa possíveis variáveis que não foram inseridas no modelo, mas que também contribuiriam para a explicação de  $Y_i$ , a relação existente entre  $Y$  e  $X$  é linear, os erros são independentes com média nula, a variância do erro é constante e o erro  $u_i, i = 1, 2, \dots, n$  são normalmente distribuídos [3]:

$$u_i \stackrel{iid}{\sim} N(0, \sigma^2)$$

## 2.3 ESTIMATIVA COM O MÉTODO MÍNIMOS QUADRADOS

O método dos mínimos quadrados consiste em adotar os estimadores que minimizam a soma dos quadrados dos desvios entre valores estimados e valores observados na amostra [7].

O objetivo é minimizar a função

$$L = \sum_{i=1}^n u_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_p x_{ip})^2$$

Derivando  $L$  em função dos  $\beta$ 's obtemos

$$\frac{\partial L}{\partial \beta_0} = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_p x_{ip})$$

$$\frac{\partial L}{\partial \beta_j} = -2 \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \dots - \beta_p x_{ip}) x_{ji} \quad j = 1, 2, \dots, p.$$

Igualando as derivadas parciais a zero e substituindo  $\beta_0, \beta_1, \dots, \beta_p$  por  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ , temos o sistema de equações:

$$\left\{ \begin{array}{l} n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_{i1} + \hat{\beta}_2 \sum_{i=1}^n x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{ip} = \sum_{i=1}^n Y_i \\ \hat{\beta}_0 \sum_{i=1}^n x_{i1} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^n x_{i1}x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{i1}x_{ip} = \sum_{i=1}^n x_{i1}Y_i \\ \vdots \\ \hat{\beta}_0 \sum_{i=1}^n x_{ip} + \hat{\beta}_1 \sum_{i=1}^n x_{i1}x_{ip} + \hat{\beta}_2 \sum_{i=1}^n x_{ip}x_{i2} + \dots + \hat{\beta}_p \sum_{i=1}^n x_{ip}^2 = \sum_{i=1}^n x_{ip}Y_i \end{array} \right.$$

Resolvendo este sistema, obtemos os estimadores de mínimos quadrados  $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$  dos parâmetros do modelo em questão.

### 2.3.1 REPRESENTAÇÃO MATRICIAL DO MRLM

Nota-se que o estimador de mínimos quadrados dos parâmetros pode ser facilmente encontrado considerando a notação matricial dos dados, que é de fácil manipulação. Desta forma, considerando a entrada de dados apresentada no modelo de Regressão Linear Múltipla pode ser escrito como [7]:

$$Y = X\beta + u_i$$

com

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad e \quad u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

em que

- $Y$  é um vetor  $n \times 1$  cujos componentes corresponde às  $n$  respostas;
- $X$  é uma matriz de dimensão  $n \times (p + 1)$  denominada matriz do modelo;
- $\epsilon$  é um vetor de dimensão  $n \times 1$  cujos componentes são os erros e
- $\beta$  é um vetor  $(p + 1) \times 1$  cujos elementos são os coeficientes de regressão.

O método de mínimos quadrados tem como objetivo encontrar o vetor  $\hat{\beta}$  que minimiza

$$L = \sum_{i=1}^n u_i^2 = u'u = (Y - X\beta)'(Y - X\beta) =$$

$$= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

sendo que  $Y'X\beta = \beta'X'Y$  pois o produto resulta em um escalar. A notação  $X'$  representa o transposto da matriz  $X$  enquanto que  $Y'$  e  $\beta'$  representam os transpostos dos vetores  $Y$  e  $\beta$ , respectivamente. Usando a técnica de derivação (em termos matriciais) obtemos

$$\frac{\partial L}{\partial \beta} = -2X'Y + 2X'X\beta$$

Igualando a zero e substituindo o vetor  $\beta$  pelo vetor  $\hat{\beta}$ , temos

$$(X'X)\hat{\beta} = X'Y$$

Em geral, a matriz  $(X'X)$  é não singular, ou seja, tem determinante diferente de zero e, portanto, possui inversa. Desta forma, conclui-se que os estimadores para os parâmetros  $\beta_j$ ,  $j = 0, \dots, p$  são dados pelo vetor

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Portanto, o modelo de regressão linear ajustado e o vetor de resíduos são respectivamente

$$\hat{Y} = X\hat{\beta} \quad \text{e} \quad u = Y - \hat{Y}$$

Ao substituir os estimadores de mínimos quadrados, tem-se que  $\hat{Y} = HY$  no qual

$$H = X(X'X)^{-1}X'$$

é a matriz chapéu, ou matriz de projeção do vetor de respostas  $Y$  no vetor de respostas ajustadas  $\hat{Y}$ .

## 2.4 MULTICOLINEARIDADE

Nos modelos de regressão linear múltipla é importante avaliar se há multicolinearidade, pois a presença de multicolinearidade pode influenciar nos dados assim ocultando algumas variáveis importantes do estudo, e tornando os parâmetros com estimativas imprecisas.

As estimativas das variâncias das estimativas dos parâmetros são indicadores adequados da existência de multicolinearidade. O efeito de uma variável explanatória pode ser suficientemente forte, de tal maneira que o respectivo coeficiente se mostre estatisticamente diferente de zero apesar da multicolinearidade; uma multicolinearidade bastante alta impedirá, entretanto, que se detecte a influência de variáveis importantes [7].

Uma análise estatística para a verificação desta multicolinearidade é a utilização do Fator de Inflação da Variância (VIF), a sua utilização facilita a verificação do grau que se encontra a relação destas variáveis.

O VIF é dado por [9]:

$$VIF_j = \frac{1}{1 - R_j^2}$$

onde,  $R_j^2$  é o coeficiente de determinação de cada variável explicativa.

O valor de aceitação da variável varia para alguns autores, para nosso modelo considere que o VIF é indicativo de problemas de multicolinearidade se  $VIF > 10$ .

## 2.5 MÉTODO STEPWISE

O método stepwise conhecido como o método do “passo a frente” ou “passo a passo”, ele ajusta o seu modelo de regressão de forma interativa a cada etapa do processo incluindo (passo forward) ou descartando (passo backward) a variável selecionada. As incorporações ou as eliminações são realizadas passo a passo até que todas as variáveis restantes sejam significativas, ou seja, entra no modelo as variáveis que possui maior coeficiente de correlação com a variável dependente do modelo. Desta forma chegamos no modelo final do processo ([8] e [11]).

Os procedimentos realizados no método stepwise são [2]:

1. Escolhe-se a variável  $x_k$  que possui o maior coeficiente de correlação para entrar no modelo.
2. Uma variável  $x_i$  entra no modelo, se o coeficiente de correlação for maior que o anterior,  $x_i$  permanece no modelo, caso contrário  $x_i$  sai do modelo.
3.  $x_i$  sai do modelo se o coeficiente de correlação for menor que o anterior,  $x_i$  fica no modelo, caso contrário,  $x_i$  permanece fora do modelo. Este passo é repetido até que não tenha mais  $x_i$  para sair do modelo. Terminada esta etapa retorna-se ao passo 2 e este passo continua até que não tenham mais variáveis para entrar no modelo.



### 3. RESULTADOS

No início do trabalho foi relatado que a mortalidade no desmame seria a principal causa do baixo resultado do LEFA. Através de análises realizadas durante o desenvolvimento do trabalho, vamos avaliar a influência das principais variáveis no resultado do processo em estudo.

Inicialmente avaliou-se as estatísticas descritivas de cada variável no processo (Tabela 3.1), assim podendo identificar inicialmente possíveis erros no banco de dados.

Tabela 3.1: Estatísticas descritivas das variáveis a serem utilizadas no estudo

Variável	Valor Máximo	Valor Mínimo	Média	Variância	Desvio Padrão
LEFA	30,64	19,58	26,12	3,76	1,94
NT	14,81	12,62	13,61	0,27	0,52
NV	13,59	11,39	12,40	0,22	0,47
NM	9,95	3,81	6,51	1,34	1,16
MM	3,66	0,97	2,34	0,34	0,58
M Desm	23,65	6,20	12,19	9,19	3,03
PFA	2,79	2,07	2,50	0,01	0,10
ID Desm	26,91	19,84	23,19	1,50	1,22
ID Creche	77,80	25,48	67,53	22,46	4,74
P Creche	25,30	5,48	19,84	4,55	2,13
M Creche	6,78	0,80	2,61	1,10	1,05
R Cio	11,74	2,09	5,48	3,69	1,92
TX Parto	98,22	79,00	90,29	8,00	2,83
M Fêmea	1,36	0	0,62	0,05	0,22
Plantel	4124	974	2620,31	682599	826,2

Na figura 3.1 são apresentadas as correlações entre as variáveis, sendo possível observar que há relações altas, moderadas, baixas e ausência de correlações nulas. A variável LEFA possui uma moderada correlação com os nascidos totais, nascidos vivos, mortalidade do desmame e mortalidade na creche, destaca-se as correlações dos nascidos totais com os nascidos vivos e a idade da creche com o peso do animal na creche. A taxa de parto também possui uma boa correlação com os nascidos totais, nascidos vivos e o peso da fêmea.

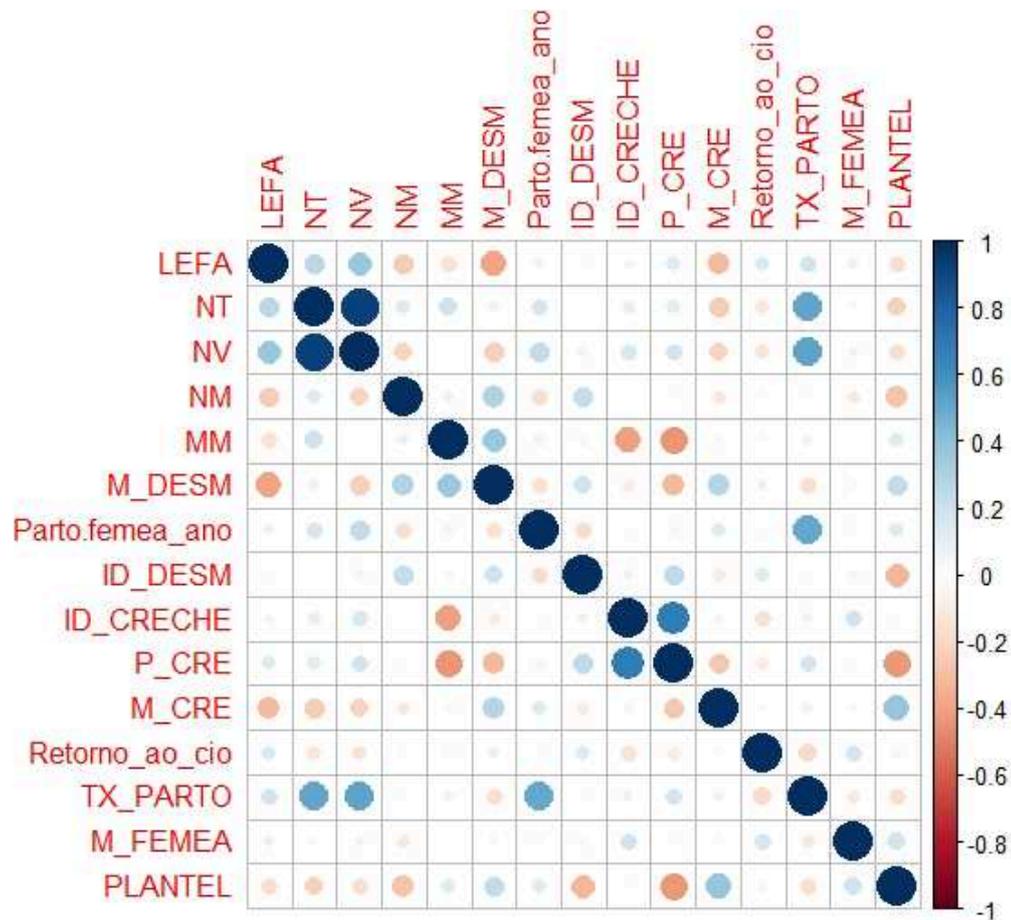


Figura 3.1: Correlação entre variáveis consideradas no estudo

Considerando a variável LEFA como a variável dependente, por ser essa a variável de interesse, toda a modelagem será visando a melhora do resultado do LEFA. As demais variáveis do estudo são as independentes.

Pelo critério VIF (menor que 10) foi selecionada as variáveis explicativas: mortos no desmame, parto fêmea ano, idade desmame, idade da creche, mortalidade da creche, retorno ao cio, taxa de parto e o plantel. Os valores do VIF de cada variável são apresentados na Tabela 3.2.

Tabela 3.2: Valores do VIF das variáveis explicativas selecionadas

Variável	VIF
Nascidos totais ( $x_1$ )	3817,47
Nascidos vivos ( $x_2$ )	3801,33
Nascidos mortos ( $x_3$ )	427,95
Mumificados ( $x_4$ )	108,01
Mortalidade no Desmame( $x_5$ )	1,68
Parto fêmea ano( $x_6$ )	1,56
Idade desmame ( $x_7$ )	1,38
Idade da creche ( $x_8$ )	3,31
Peso da creche ( $x_9$ )	3,87
Mortalidade da creche ( $x_{10}$ )	1,43
Retorno ao cio ( $x_{11}$ )	1,17
Taxa de parto ( $x_{12}$ )	2,05
Mortalidade fêmea ( $x_{13}$ )	1,19
Plantel ( $x_{14}$ )	2,13

O próximo passo consiste em fazer a seleção variáveis para o modelo final via critério stepwise, considerando as explicativas selecionadas com VIF menor que 10. As variáveis que permaneceram no modelo final foram: taxa de parto, retorno ao cio, mortalidade da creche e mortos no desmame. Na tabela 3.3 são apresentadas as estimativas do modelo final.

Tabela 3.3: Análise de regressão das variáveis explicativas selecionadas

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	18,96002	491,256	3,860	0,000172
Mortalidade Desmame	-0,20506	0,04953	-4,140	5,93e-05
Mortalidade Creche	-0,37039	0,14138	-2,620	0,009754
Retorno ao cio	0,19150	0,07544	2,538	0,012214
Taxa Parto	0,10599	0,05199	2,039	0,043340

Pelo ajuste do modelo final (conforme tabela 3.3), todas as variáveis são significativas ao nível de 5% de significância. Abaixo segue pelo modelo da regressão linear múltipla estimada os coeficientes que impacta no resultado do LEFA.

$$Y = 18,96 - 0,205x_5 - 0,37x_{10} + 0,192x_{11} + 0,106x_{12}$$

Desta forma pelo modelo de previsão, para melhorar o resultado será necessário a redução de mortalidade ( $x_5$  e  $x_{10}$ ) ou aumentar a taxa de parto ( $x_{12}$ ) e retorno ao cio ( $x_{11}$ ).

Nas figuras 3.2 a 3.5 são apresentadas as análises de resíduos do modelo final. A figura 3.2 representa a homoscedasticidade dos resíduos, indica se as variáveis possui linearidade com a variável resposta do modelo. No gráfico os valores estão bem distribuídos em relação a linha 0 e não indica linearidade nos dados.

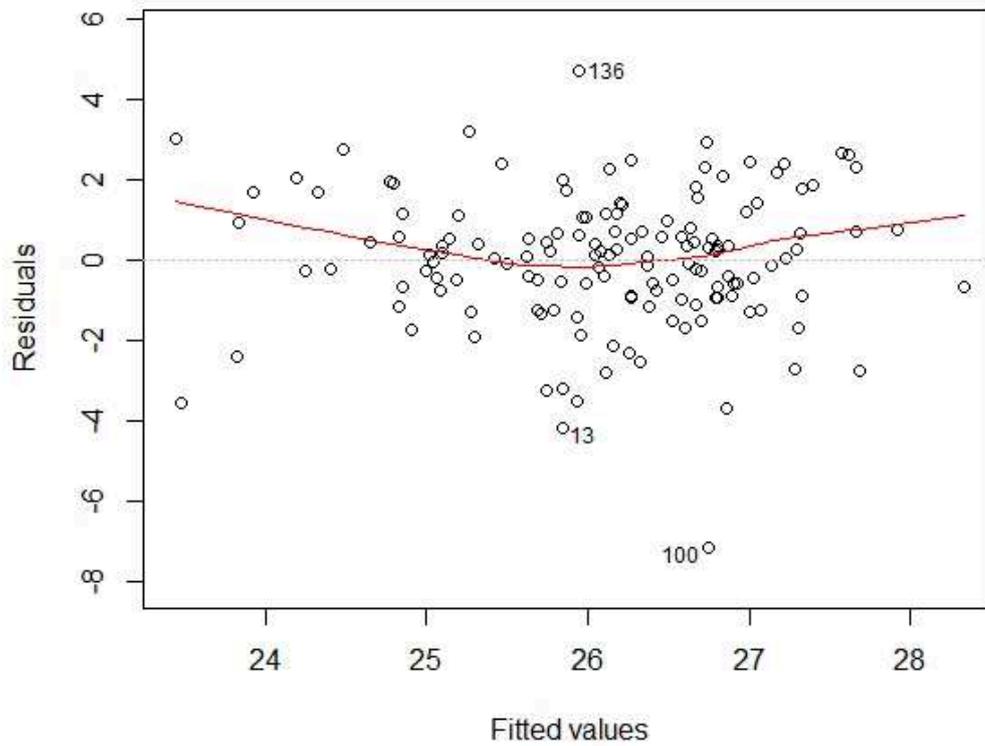


Figura 3.2: Homoscedasticidade dos Resíduos

A figura 3.3 representa a normalidade dos resíduos. O gráfico indica normalidade dos resíduos, pode-se dizer que os resíduos possuem distribuição normal.

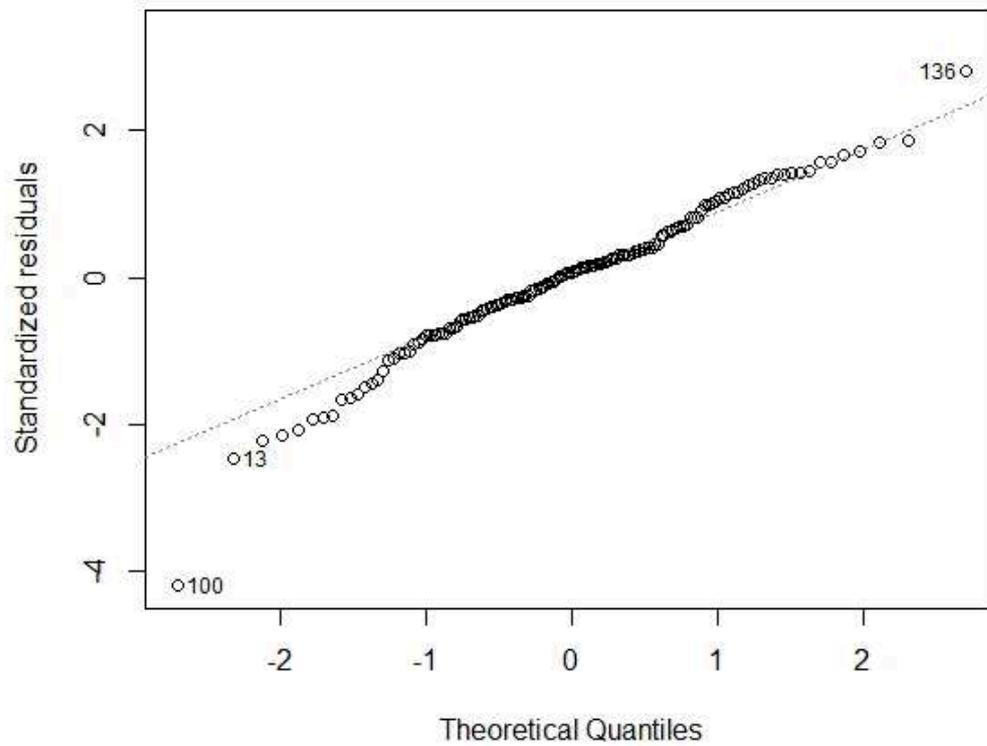


Figura 3.3: Normalidade dos Resíduos

A figura 3.4 apresenta o gráfico de escala, este gráfico mostra se os resíduos estão espalhados igualmente ao longo do intervalo de previsão. O ideal é que a linha neste gráfico seja horizontal, com pontos aleatoriamente espalhados no gráfico. O gráfico abaixo apresenta uma leve variação nos dados ao meio do intervalo, portanto pode-se dizer que possui uma variação uniforme ao longo do intervalo.

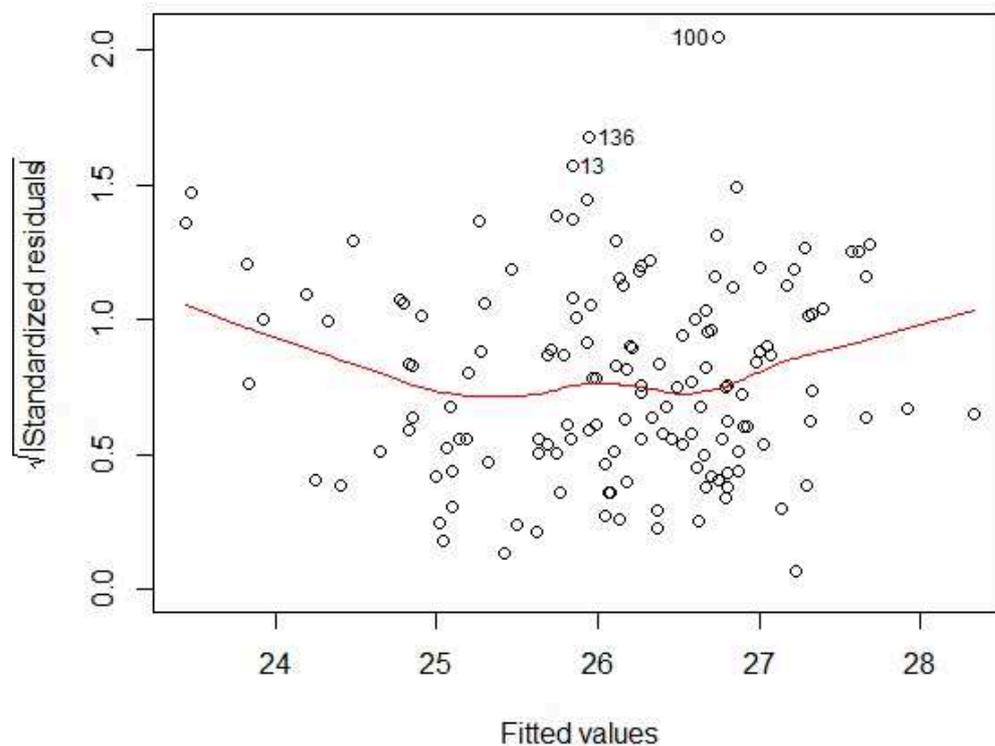


Figura 3.4: Distribuição dos Resíduos

A figura 3.5 ajuda a identificar dados que podem influenciar no ajuste do modelo, ou seja, os outliers. O gráfico possui uma distancia de Cook de 0,5 (representado pela linha pontilhada), se houver pontos fora deste intervalo é indicio de ser um outlier, portanto devem ser removidos gerando um novo ajuste para a regressão. Por este gráfico não existe nenhum ponto de influência no modelo, não sendo necessário a retirada de nenhum ponto do modelo.

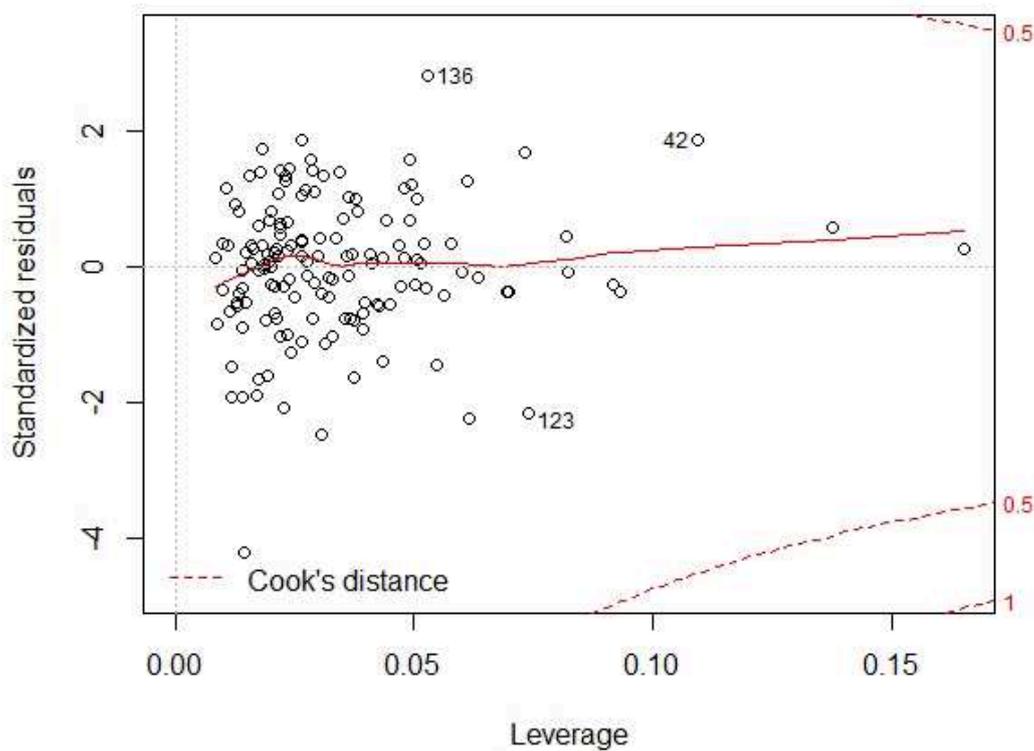


Figura 3.5: Pontos Influentes no modelo

Observe abaixo as condições de ajuste com redução ou aumento de 5 a 10% na média de uma variável e mantendo o resultado médio das demais variáveis. Segue as médias das variáveis obtidas na estatística descritiva:

LEFA(Y)=26,12;

Mortalidade do desmame( $x_5$ ) = 12, 19;

Mortalidade na creche( $x_{10}$ ) = 2, 61;

Retorno ao cio( $x_{11}$ ) = 5, 48;

Taxa de parto( $x_{12}$ ) = 90, 29;

*Caso1* : Considerando a redução de 10% na mortalidade do desmame( $x_5 = 10, 971$ ) e mantendo os resultados médios da mortalidade na creche( $x_{10}$ ), retorno ao cio ( $x_{11}$ ) e taxa de parto( $x_{12}$ ), obtêm-se o valor esperado do LEFA de 26, 37.

*Caso2* : Considerando a redução de 10% na mortalidade da creche( $x_{10} = 2, 349$ ) e mantendo os resultados da mortalidade do desmame( $x_5$ ), retorno ao cio( $x_{11}$ ) e taxa de parto( $x_{12}$ ), obtêm-se o valor esperado do LEFA de 26, 21

*Caso3* : Considerando o aumento de 10% no retorno ao cio ( $x_{11} = 6,028$ ) e mantendo os resultados da mortalidade do desmame ( $x_5$ ), mortalidade na creche ( $x_{10}$ ), e taxa de parto ( $x_{12}$ ), obtêm-se o valor esperado do LEFA de 26,22

*Caso4* : Considerando o aumento de 5% na taxa de parto ( $x_{12} = 94,804$ ) e mantendo os resultados da mortalidade do desmame ( $x_5$ ), mortalidade na creche ( $x_{10}$ ) e retorno ao cio ( $x_{11}$ ), obtêm-se o valor esperado do LEFA de 26,59

*Caso5* : Para a entrega do LEFA na meta desejada, ou seja, LEFA igual a 26,8 e considerando que irá atuar na mortalidade no desmame (segundo maior impacto, conforme simulado acima), o valor desta variável deverá ser 8,86, portanto, uma redução de 27,32% na média atual da mortalidade no desmame.

Neste caso para chegar no resultado do LEFA de 26,8 precisa reduzir a mortalidade em 27,32%. Analisando os dados a melhor variável com um impacto significativo é a taxa de parto e a mortalidade do desmame.

Pode-se avaliar este impacto dos indicadores em valores financeiros, o LEFA representa o número médio de leitões entregues, o valor de 1 leitão entregue a mais neste indicador representa R\$240.000,00/mês. Apresentando na tabela 3.4 a evolução do indicador LEFA, lembrando que a média atual é de 26,12 e que a cada aumento de 1,0 no LEFA representa os R\$240.000,00/mês.

Tabela 3.4: Impactos financeiros, mensais e anuais, de acordo com a evolução do indicador LEFA

LEFA	Impacto indicador	Impacto Mês (MR\$)	Impacto Ano (MR\$)
26,12	0,00	0,00	0,00
26,25	0,13	31,2	374,4
26,39	0,27	64,8	777,6
26,75	0,63	151,2	1.814,4
27,00	0,88	211,2	2.534,4
27,12	1,00	240	2.880

MR\$: milhares de reais

Para a melhora na mortalidade do desmame os produtores terá que realizar um bom manejo (rotatividade) dos leitões recém nascidos fazendo que todos os animais se amamentem o tempo mínimo e garanta o rodizio na primeira mamada pois o colostro traz imunidades para o leitão. Manter a área de maternidade limpa, desinfetada e com temperatura adequada para evitar as doenças e bactérias.

A Mortalidade de creche possui impactos por doenças ou bactérias, estas doenças são provenientes dos leitões que tiveram falhas sanitárias e de manejo na sua fase de desmame, ou de animais que estavam alojados anteriormente no local, portanto deve realizar o vazio sanitário entre um alojamento e outro de no mínimo 5 dias para a desinfecção do crechário.

No retorno ao cio os produtores devem padronizar o processo de inseminação e após o seu ultimo período de gestação dar um intervalo mínimo para a recuperação da fêmea até a nova inseminação.

Na taxa de parto sugere-se melhorar o local onde as matrizes ficam alojadas durante o seu período de gestação, o ambiente deve possuir uma temperatura adequada e constante, um local

limpo para evitar contaminação, um bom espaço e redução de ruídos para evitar o estresse e contaminações que causam os abortos.



## 4. CONCLUSÕES

Com este trabalho foi possível concluir que através de análises estatísticas temos quatro variáveis principais que impactam o resultado do LEFA, destas variáveis temos a mortalidade no desmame e a mortalidade na creche que pioram o nosso resultado, ambos possuem uma relação negativa, portanto são variáveis que precisam ter redução no percentual de seus resultados, melhorando as condições sanitárias dos crechários, realizando um bom manejo dos animais, ajustando temperaturas do ambiente para reduzir stress e proliferação de bactérias.

As outras duas variáveis encontradas que impactam no resultado do indicador é o retorno ao cio e a taxa de parto, essas variáveis possuem relação positiva, portanto contribuem positivamente ao resultado do LEFA, quanto maior melhor, portanto a taxa de parto precisa ter seu resultado melhorado com as melhoras nas condições sanitárias e no ambiente que os animais ficam durante o seu período de gestação. Trabalhando com estas quatro variáveis principais pode-se chegar ao resultado ideal do LEFA que é 26,8.

As demais variáveis explicativas do modelo não foram significativas, portanto ao nível de 5% de significância, estas variáveis não influenciam no resultado do processo.



# REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ABPA: *Associação Brasileira De Proteína Animal*, 2018. <http://abpa-br.com.br/>, acessado em 15/06/2019.
- [2] Alves, M. F., Lotufo, A. D. P. e Lopes, M. L. M.: *Seleção de variáveis stepwise aplicadas em redes neurais artificiais para previsão de demanda de cargas elétricas*. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, 1(1), 2013.
- [3] Charnet, R., Freire, C. d. L., Charnet, E. M. R., Bonvino, H. *et al.*: *Análise de modelos de regressão linear com aplicações*. Campinas, São Paulo, Unicamp, 356p, 1999.
- [4] Fávero, L. P. L., Belfiore, P. P., Silva, F. L. d. e Chan, B. L.: *Análise de dados: modelagem multivariada para tomada de decisões*. 2009.
- [5] Guimarães, T. P., Silva, M. A. P. da e Leão, K. M.: *Índices zootécnicos de uma granja produtora de leitões*. PUBVET, 4:Art-980, 2010.
- [6] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. e Tatham, R. L.: *Análise multivariada de dados*. Bookman Editora, 2009.
- [7] Hoffmann, R. e Vieira, S.: *Análise de regressão: uma introdução à econometria*. São Paulo, 1998.
- [8] Keppel, G.: *Design and analysis: A researcher's handbook*. Prentice-Hall, Inc, 1991.
- [9] Maroco, J.: *Análise estatística com utilização do SPSS.(2007)*. Lisboa: Edições Sílabo, 3.
- [10] Silva, E. R.: *SISTEMA PRODUTOR DE LEITÕES-SPL*.
- [11] Tabachnick, B. e Fidell, L.: *Using Multivariate Statistics (5. ° ed.)*, 1996.
- [12] Tenfen, L. W.: *Efeito do peso ao nascimento de leitões sobre o ganho médio diário e peso ao desmame nas variadas ordens de parto*. Medicina Veterinária-Tubarão, 2017.