



Regressão linear

Transcrição

Ao analisarmos o gráfico de vendas de brinquedo por mês, percebemos uma tendência de crescimento. Além da reta de crescimento, temos uma **sazonalidade**. Temos alguns picos de maior venda e outros de baixa. Também existem alguns fatores que não conseguimos identificar através dos dados. Isto em estatística se chama **erro**, a parte da variação que não conseguimos explicar.

Em seguida, vamos decompor o gráfico em três módulos: **tendência**, **sazonalidade** e **erro**. Como faremos para demonstrar estes pontos em uma reta? Quem já fez o [curso de estatística \(https://www.alura.com.br/course-online-introducao-a-estatistica-1\)](https://www.alura.com.br/course-online-introducao-a-estatistica-1) da Alura conhece **regressão linear**. Trata-se de uma equação para se estimar a condicional (valor esperado) de uma variável y , dados os valores de algumas outras variáveis x . Demonstraremos visualmente adiante.

É possível fazê-la manualmente, mas o Google Spreadsheets (assim como o Excel) já tem essa função. O ponto y da reta é um fator que chamamos de m (o coeficiente angular da reta) multiplicado por x , em seguida, somado com b (o fator linear).

$$y = (m * x) + b$$

[COPIAR CÓDIGO](#)

O x seria referente a data. Seria complicado usar uma data, então criaremos uma coluna auxiliar: $Aux - x$. Os valores serão triviais.

Mês	Brinquedos Para Gatos	Aux - x
1-Jan-2013	13	1
1-Feb-2013	19	2
1-Mar-2013	30	3
1-Apr-2013	25	4
1-May-2013	31	5
1-Jun-2013	37	6
1-Jul-2013	63	7
1-Aug-2013	56	8
1-Sep-2013	53	9
1-Oct-2013	59	10
1-Nov-2013	50	11
1-Dec-2013	98	12
1-Jan-2014	73	13
1-Feb-2014	92	14
1-Mar-2014	108	15
1-Apr-2014	71	16
1-May-2014	82	17
1-Jun-2014	98	18
1-Jul-2014	154	19
1-Aug-2014	130	20
1-Sep-2014	112	21
1-Oct-2014	117	22
1-Nov-2014	101	23

A coluna Brinquedos Para Gatos será o y e Aux - x será x . Para fazer o cálculo usaremos a função `LINEST`. Ela irá pedir os dados conhecido em y e depois se queremos calcular o b ou outras informações extras. Iremos indicar as células e indicar que queremos calcular b , com `true`. Não queremos outras informações.

`=LINEST(B2:B37; D2:D37; True; False)`

COPIAR CÓDIGO

Ele irá nos devolver os resultados de m e b , respectivamente: 5,589 e 7,176. Podemos agora usar a fórmula da regressão linear. Adicionaremos um cifrão, porque os valores valerão para a reta inteira.

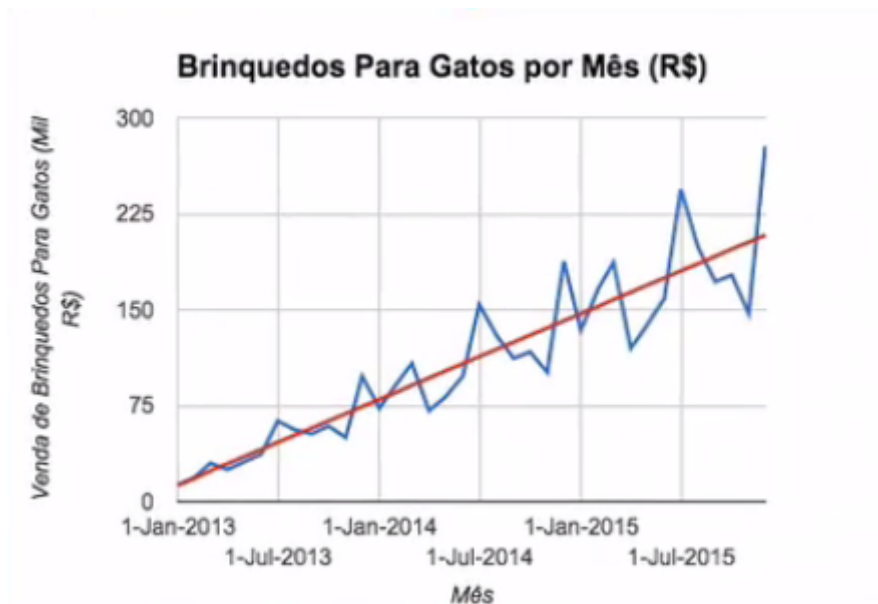
	Aux - x
$=($F$8*D2)+$G8	1
	2
	3

Vamos desenhá-la no gráfico. Vamos até configurações avançadas e estenderemos os valores selecionados para a terceira coluna.

Chart Editor



O nosso gráfico ficará assim:



Temos os dados reais e uma reta ajustada que me dá o meu crescimento.

