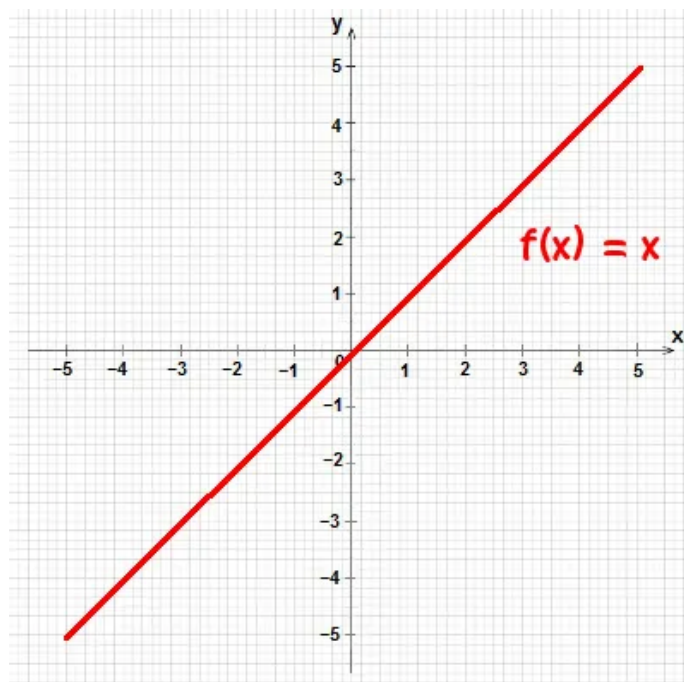


## O que Regressão Linear tem a ver com Dinossauros

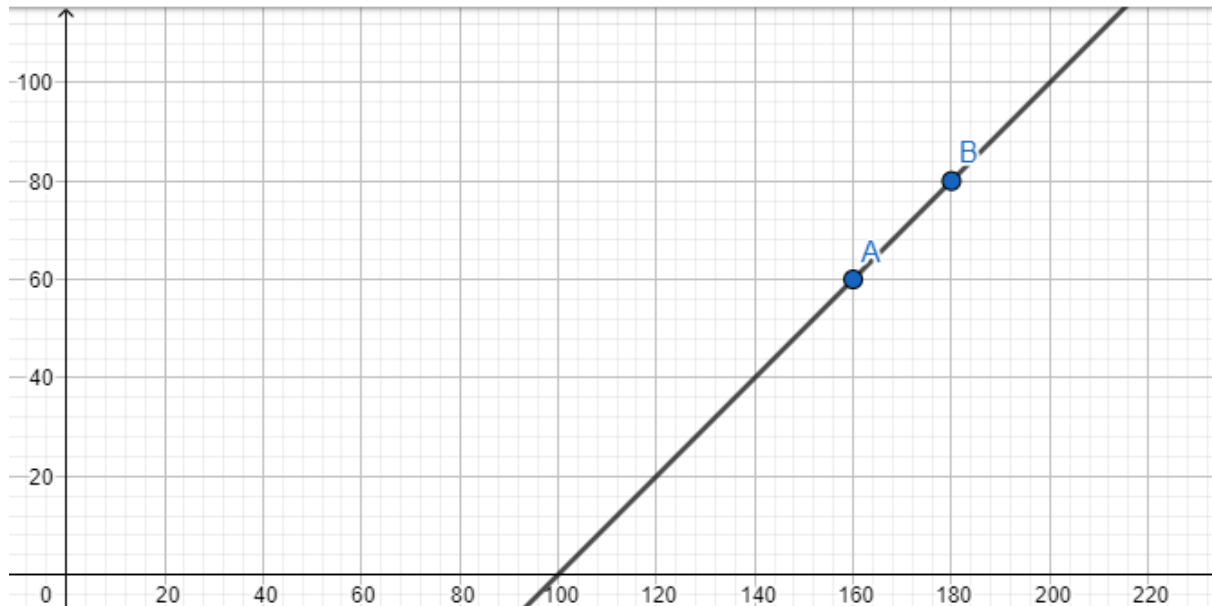


Lembra na escola quando você estava começando a aprender sobre gráficos e tinha aquela reta  $y = ax + b$  que parecia mais ou menos com isso:



Nessa época, a tia Fefê, não você já era bem grandinho nessa época, o Professor Cleidson pedia para você descobrir o valor de um determinado  $y$  olhando para a posição de um determinado número  $x$  no eixo "deitado", aí você subia o dedo até a reta e via a "altura" desse ponto, o  $y$ .

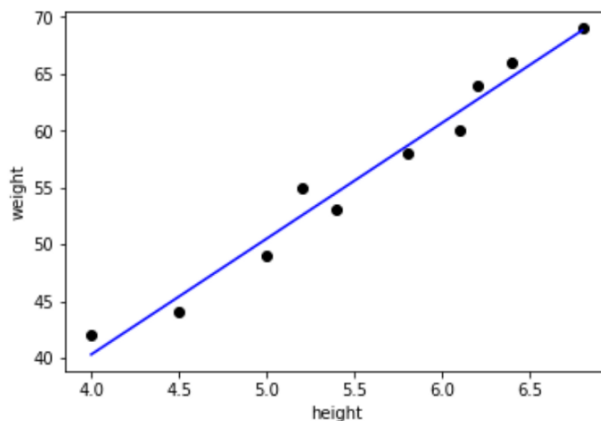
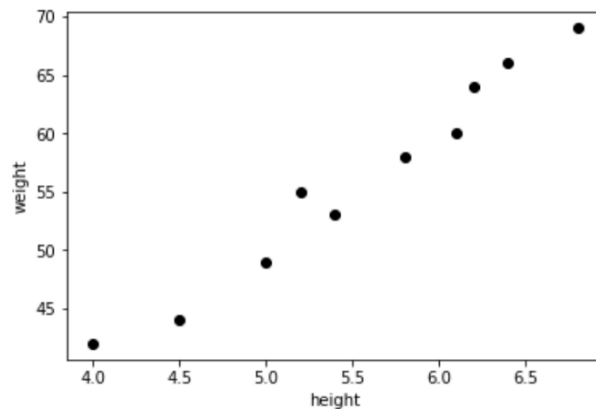
Por exemplo, no gráfico imaginário abaixo (tamanho  $x$  peso), uma pessoa de 160 cm de altura tem a massa de 60 Kg.



Um processo muito simples não? Se na vida real fosse assim e você quisesse saber seu peso se tivesse 20 metros de altura também seria muito simples, só olhar o  $x=2000$  cm subir o dedo até a reta e ver o  $y$  (Você pesaria 1904 Kg segundo esse gráfico inventado).

O problema é que na vida real as coisas não são assim tão simples e você não pode simplesmente fazer um gráfico com uma reta que magicamente prevê o peso. Mas seria muito útil se pudéssemos fazer algo que, embora não seja 100% preciso, nos dá uma boa noção, certo?

É para isso que serve a Regressão Linear. É basicamente uma forma de achar uma reta que tenha a menor distância dos pontos possível e portanto, que seja possível supor que um valor bem distante  $x$  tenha um correspondente  $y$  próximo ao estimado pela "linha".

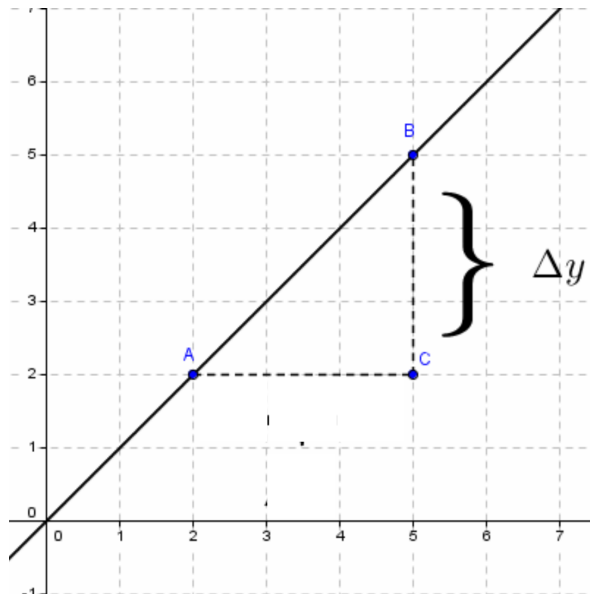


Como achar essa reta?

Você pode pegar vários valores de medições na vida real, por exemplo o peso e altura de pessoas reais, colocar no gráfico e chutar várias retas(chutar coeficientes angulares e interceptos pra quem já entende mais do assunto) e medir a distância dos pontos até ela, quanto menor a diferença, melhor.



Como você mediria a distância de um ponto a uma reta? Você trava o valor de  $x$  do ponto e mede a diferença de  $y$  com o  $\hat{y}$  da reta, ou seja  $y - \hat{y}$ . Você só faz esse processo para todos os pontos e soma tudo, você terá a soma das distâncias, aí é só achar a reta que tenha a menor soma das distâncias, certo?



Exemplo, queremos saber o  $y - \hat{y}$  de C, travamos o x no 5 subtraímos 2 de 5 (Altura de B) e achamos a distância  $y - \hat{y} = -3$ . O C está abaixo da reta por isso deu negativo.

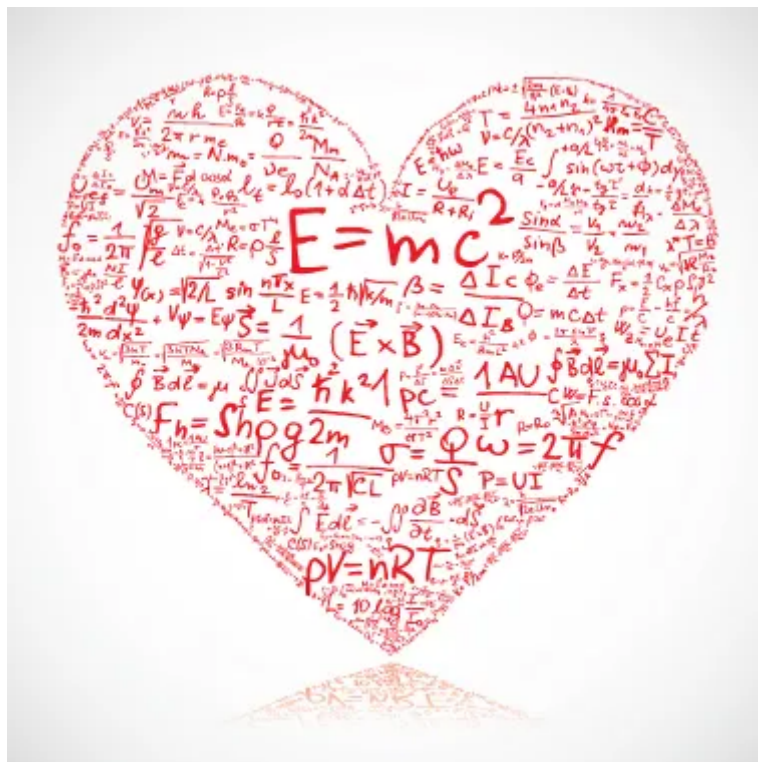
ERRADO, se existe um ponto acima da reta, o  $y$  é maior que o  $\hat{y}$  da reta e  $y - \hat{y}$  fica positivo, mas se  $y$  é menor que  $\hat{y}$ , a diferença fica negativa e se você somar a distância dos pontos a reta, os valores negativos anulam os positivos inviabilizando nosso processo genial.

Ainda bem que no mundo já viveram pessoas muito inteligentes e um cara chamado Laplace teve a ideia de somar não a diferença dos pontos as retas ( $y - \hat{y}$ ), mas sim o quadrado dessas distâncias  $((y - \hat{y})^2)$ .



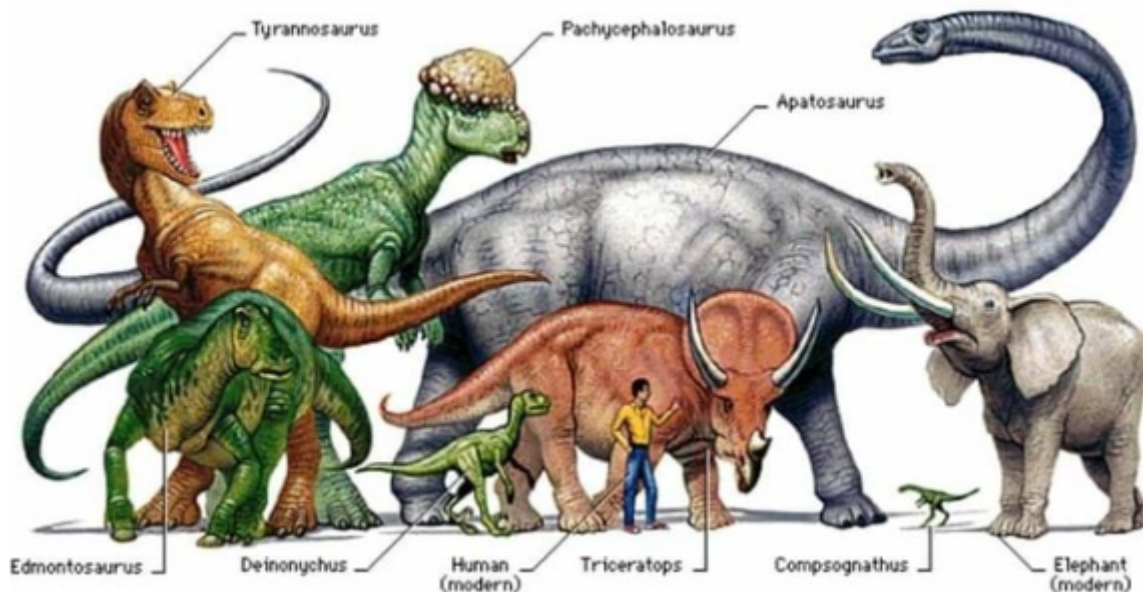
Laplace

Existem outras vantagens para se usar a distância ao quadrado, mas ei, isso aqui é só uma introdução ao assunto! Aliás, os matemáticos já criaram técnicas para evitar ficar testando várias e várias possíveis retas, mas isso requer um grande preparo de matemática universitária.



Acontece que mensurar o peso de um animal gigantesco como um T-rex não é simples, pois não existem bichos terrestres hoje que tenham tamanho próximo ao que era um grande dinossauro. O elefante é do tamanho de um triceratops, mas ele nem era lá um dos maiores.

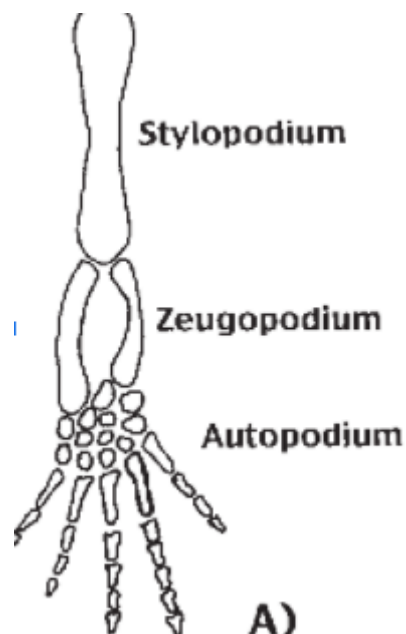




Eu queria ter 10% da confiança desse cara na imagem

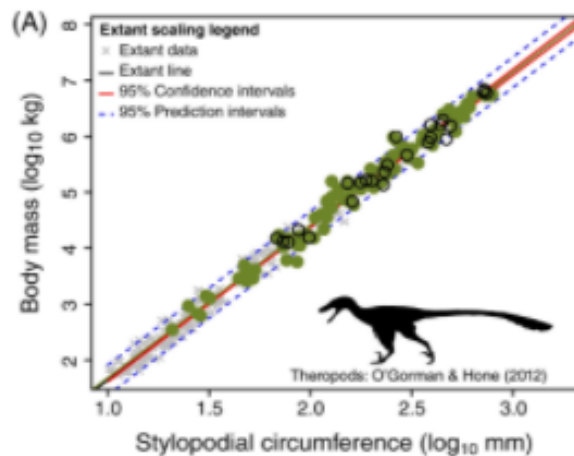
Então, para estimar o tamanho dos dinossauros, os cientistas analisaram várias qualidades de bichos vivos hoje e viram qual delas está mais relacionada com o peso.

Embora o tamanho do animal previsse bem o peso, descobriram que a circunferência de um ossinho da perna chamado stylopodium é bem melhor para fazer essa previsão.

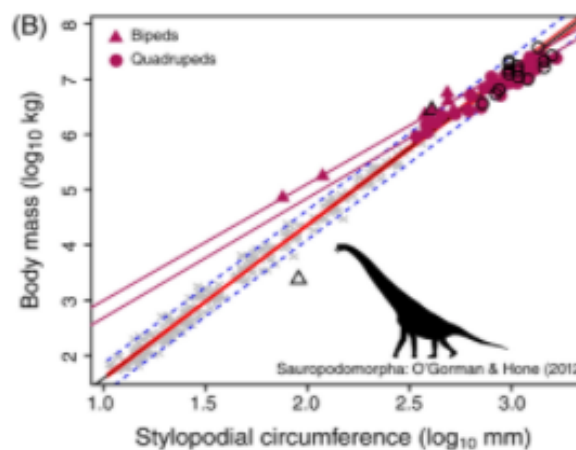


Então os paleontólogos analisaram a circunferência desses ossinhos e o peso de vários animais de hoje, fizeram uma Regressão Linear e acharam uma reta que consegue prever o peso de bichos bem melhor que em métodos que usávamos antigamente!

Eles mediram o tamanho desse osso em fósseis de dinossauros e assim descobriram o peso absurdo desses colossos!



Os terópodes, grupo do T-Rex, variavam muito em massa, mas chegavam a um peso na base de 1.000.000 Kg, Milhares de toneladas. O T-Rex pesava até 7 toneladas.



Os Saurópodes, os famosos pescoçudos, eram os maiores dinossauros, alguns deles pesavam até 100.000.000 kg, o Argentinosaurus pesava entre 50 e 100 toneladas, um ônibus pesado tem 14 toneladas, é mole?!

O que é Regressão Linear:

Por Fernando Matsumoto, Guilherme Fernandes e Piero Esposito, [Modelos de Predição | Regressão Linear | by Piero Esposito | Turing Talks | Medium](#)

Estatística: O que é, para que serve, como funciona Livro por Charles Wheelan

Estudo de Nicolás E. Campione que fala do método de medir o osso stylopodium e de onde tirei as imagens: [The accuracy and precision of body mass estimation in non-avian dinosaurs - Campione - 2020 - Biological Reviews - Wiley Online Library](#)

Escrito por Kevin Padian, peso do T-rex e ele é um terópode: [Tyrannosaur | dinosaur group | Britannica](#)

Por John P. Rafferty, peso Argentinossauro e ele é um Saurópode: [Titanosaurs: 8 of the World's Biggest Dinosaurs | Britannica](#)

Segundo a Agência Reguladora de Serviços Públicos Delegados de Transporte do Estado de São Paulo, o peso de um ônibus: [Diário das leis - Aprova as Especificações Técnicas de Veículos Automotores de Transporte Coletivo de Passageiros Rodoviário e Urbano Intermunicipal \(diariodasleis.com.br\)](#)

Laplace criou soma do quadrado das diferenças: Stigler (1986, p 153) Analysis of variance

Imagens em ordem

1º <https://82.146.63.105/en/create/template/881018>

2º [Função Linear. Função Linear: um tipo especial de função afim \(uol.com.br\)](#)

3º Feito em <https://www.geogebra.org/>

4º e 5º [Modelos de Predição | Regressão Linear | by Piero Esposito | Turing Talks | Medium](#)

6º [Roll Safe, the Guy-Tapping-Head Meme, Explained \(nymag.com\)](#)

7º [graphics - How can I use LaTeX to mark my delta y and my delta x? - TeX - LaTeX Stack Exchange](#)

8º [Opera Mundi: Hoje na História: 1827 – Morre o físico e matemático francês Pierre-Simon Laplace \(uol.com.br\)](#)

9º [A equação do amor - Escola Kids \(uol.com.br\)](#)

10º [\(3\) Pinterest](#)

11º [Figure 1 from The tetrapod limb: a hypothesis on its origin. | Semantic Scholar](#)

12º e 13º [The accuracy and precision of body mass estimation in non-avian dinosaurs - Campione - 2020 - Biological Reviews - Wiley Online Library](#)

Obs: Disseram no discord que a bibliografia não é considerada na contagem de caracteres, estou contando com isso kkkkkj, obrigado!

Obs2: Sou calouro da estatística e não tive essa matéria ainda, desculpa se errei em algum detalhe.

Por João Pedro Apolonio de Sousa Matos NUSP: 12558360