Data Science Communication

Chelsea Parlett-Pelleriti

Is this Kind?

Is this Clear?

Is this Simple?

Is this Relatable?

Is this **Kind**?

Is this **Clear**?

Is this **Simple**?

Is this **Relatable**?

David Robinson
@drob

Communication tip: When you're writing for an audience of varying experience levels, explain concepts using language that to experts doesn't feel like an explanation

Not enough detail for beginners:

We used logistic regression.

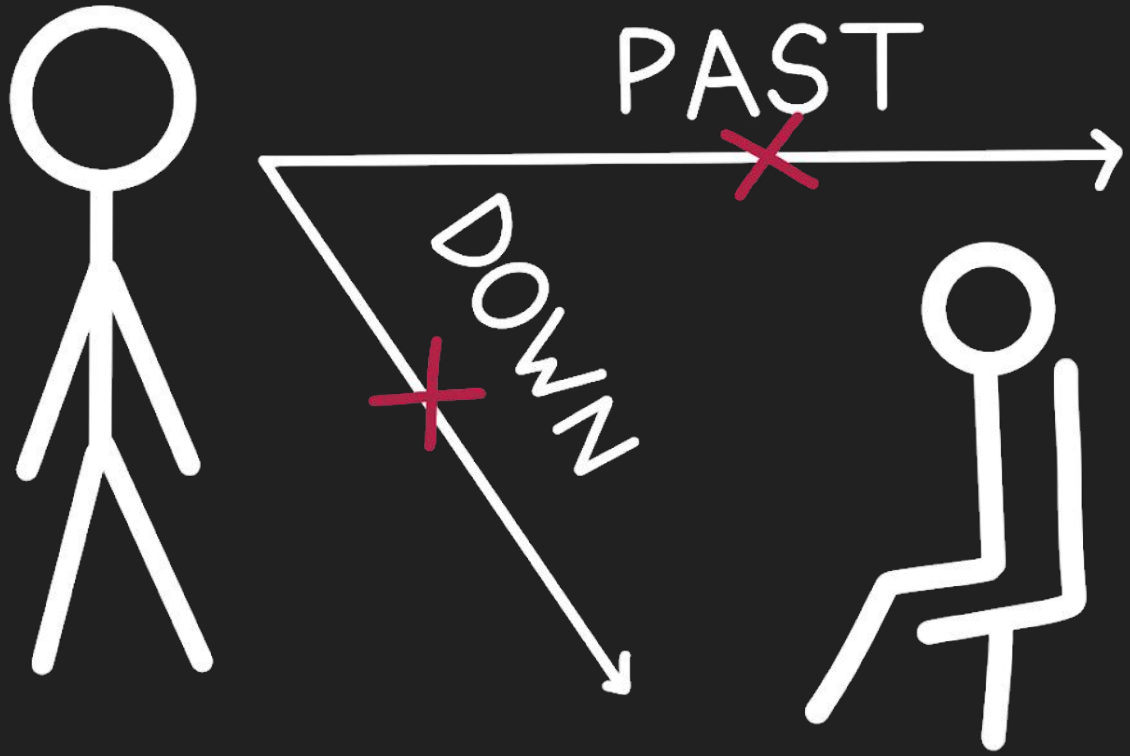
Too didactic for experts:

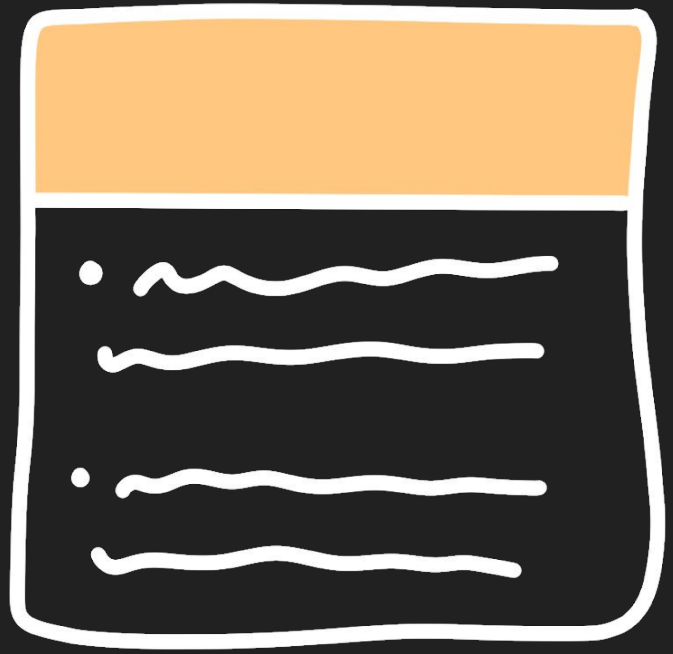
We used logistic regression. Logistic regression is a model suited for predicting a binary (“yes” or “no”) outcome.

Readable for everyone:

Since our outcome is binary (“yes” or “no”), we used logistic regression.

8:44 AM · Aug 24, 2020 · Twitter Web App


PAST
DOWN



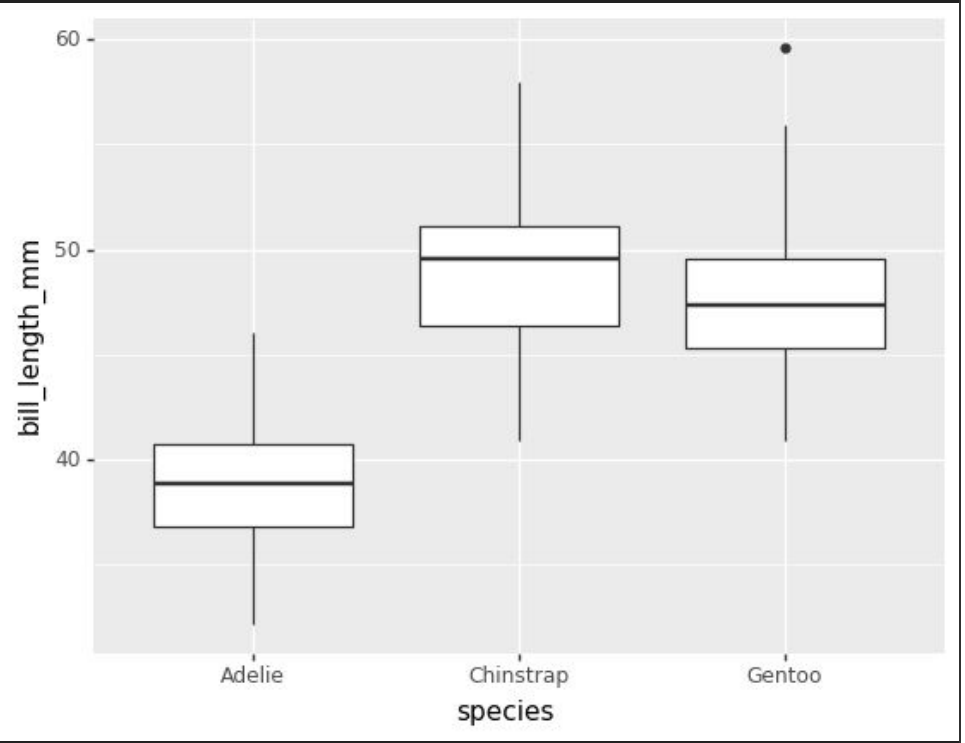
What's a time you felt like someone **didn't** have empathy when they explained something to you? How did that make you feel?

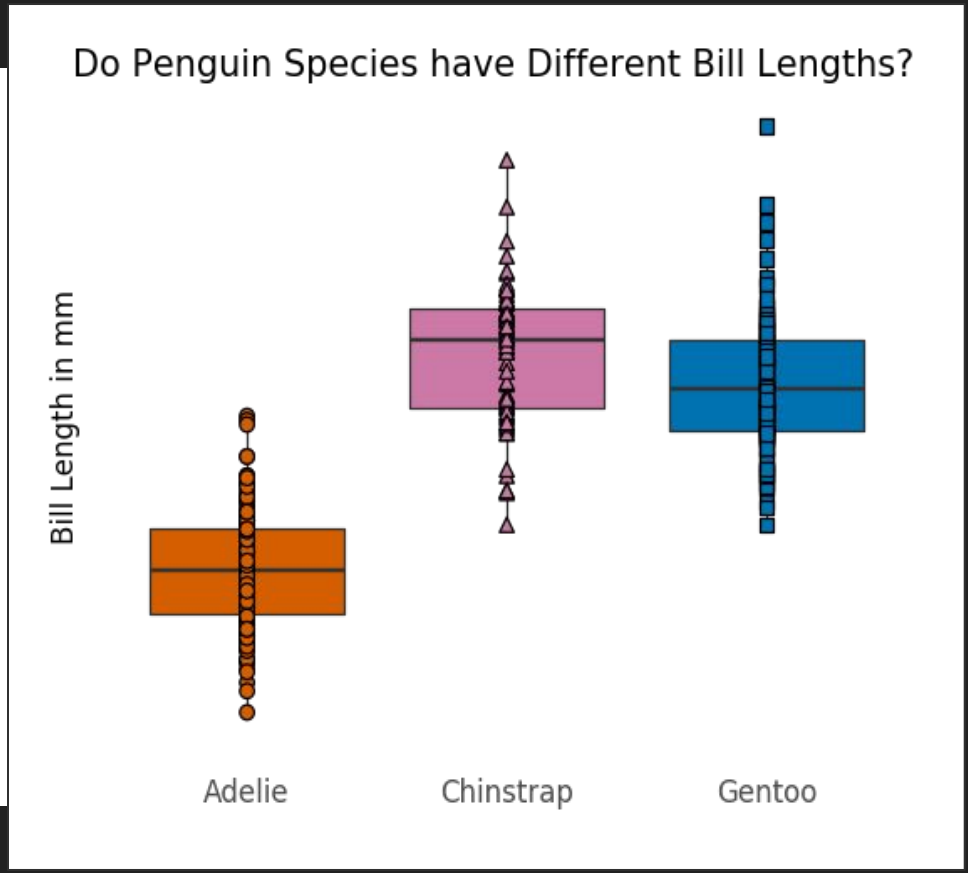
Is this Kind?

Is this Clear?

Is this Simple?

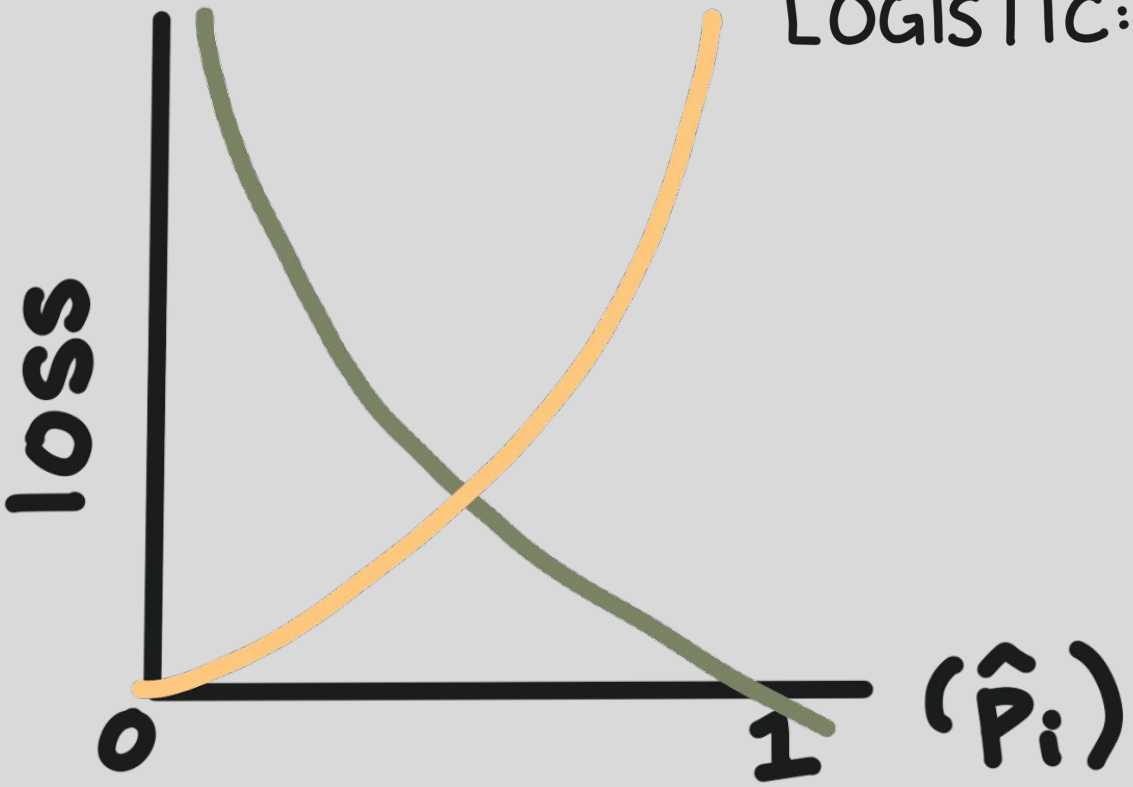
Is this Relatable?


60
50
40
bill_length_mm
Adelie
Chinstrap
Gentoo
species


Do Penguin Species have Different Bill Lengths?
Bill Length in mm
Adelie
Chinstrap
Gentoo

LINEAR: $(\hat{y}_i - y_i)^2$

LOGISTIC:
$$\begin{cases} -\log(\hat{p}_i) & \text{if } y=1 \\ -\log(1-\hat{p}_i) & \text{if } y=0 \end{cases}$$



What's one thing that you've noticed someone (a teacher, someone on social media...) do to **make their message clear**?

Is this Kind?

Is this Clear?

Is this Simple?

Is this Relatable?

Simplify

xkcd.com/simplewriter/

Positive Correlation: when two things move in the same direction in a very close way.

Loss Function: a number that tells you if something is doing a good job, we want this number to be as small as possible. Small means you are doing a good job. Big means you are doing a bad job.

Homoscedasticity: Things will be spread out a bit. But we want them to be spread out in the same way, no matter where you look.

Bias: When you're wrong in the same way over and over.

Regression

Regularization

Bootstrapping

Gradient

P-value

Posterior

Confidence Interval

A/B test

Approximate Methods

Cohesion and Separation

Convolution

Median

Neural Network

Standard Deviation

Variance

Vectorization

Convergence

Newton's Method

Recursion

Statistical Power

Decision Tree

Effect Size

Data Frame

Is this Kind?

Is this Clear?

Is this Simple?

Is this Relatable?

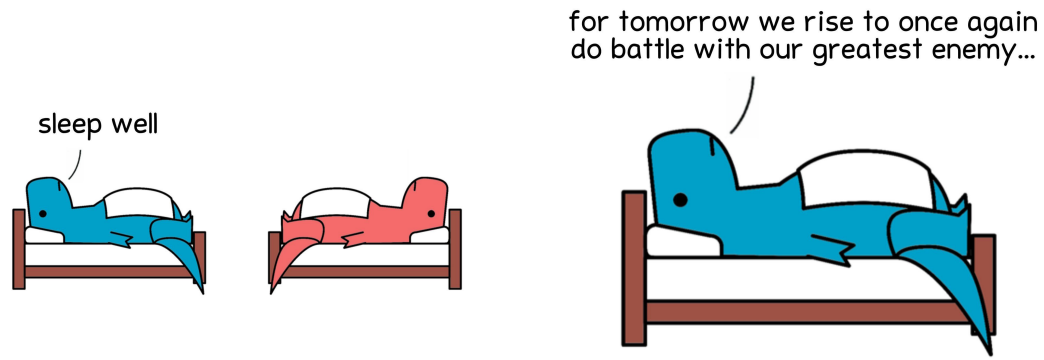

TikTok

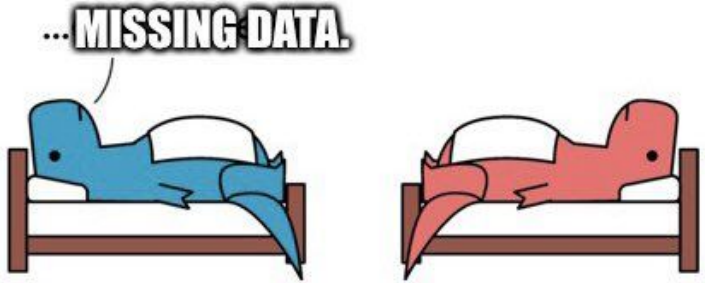

LASSO
SMALL
BUT NON-ZERO
COEFFICIENTS


ME PROUD
OF 99%
ACCURACY


IMBALANCED
CLASSES


LINEAR REGRESSION
LOG ODDS
VIA 9GAG.COM
LOGISTIC REGRESSION
imgflip.com

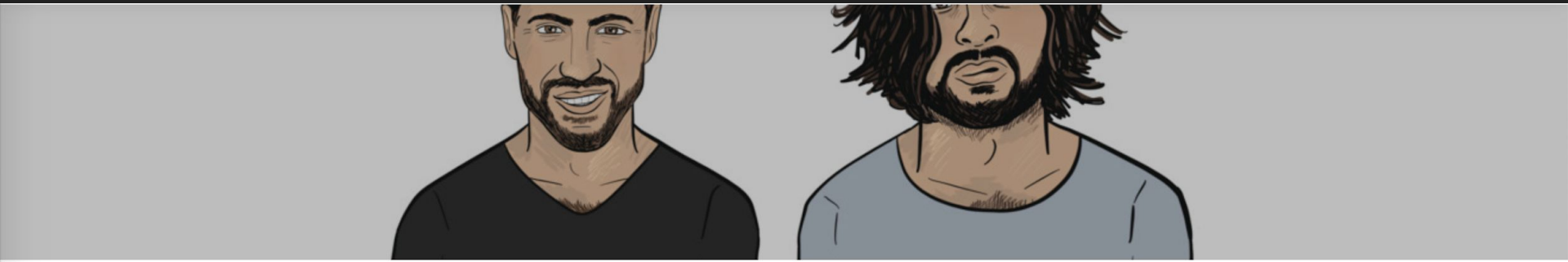

sleep well
for tomorrow we rise to once again do battle with our greatest enemy...


...MISSING DATA.

Me explaining why standardizing your variables is important:

#RStats

Decided to put coin in a bucket everytime I FORGET GGPLOT SYNTAX AND HAVE TO LOOK IT UP ... Started 1 hour ago



107: Memes, TikTok, and science communication (with Chelsea Parlett-Pelleriti)

Episode 107 · May 4th, 2020 · 1 hr 5 mins

Is this Kind?

Is this Clear?

Is this Simple?

Is this Relatable?