



Hackademy 2015 - Data Analytics

Dan Couture
@MathYourLife

May 15, 2015

INTERNET
PERFORMANCE.
DELIVERED.

 dyn.com  @dyn

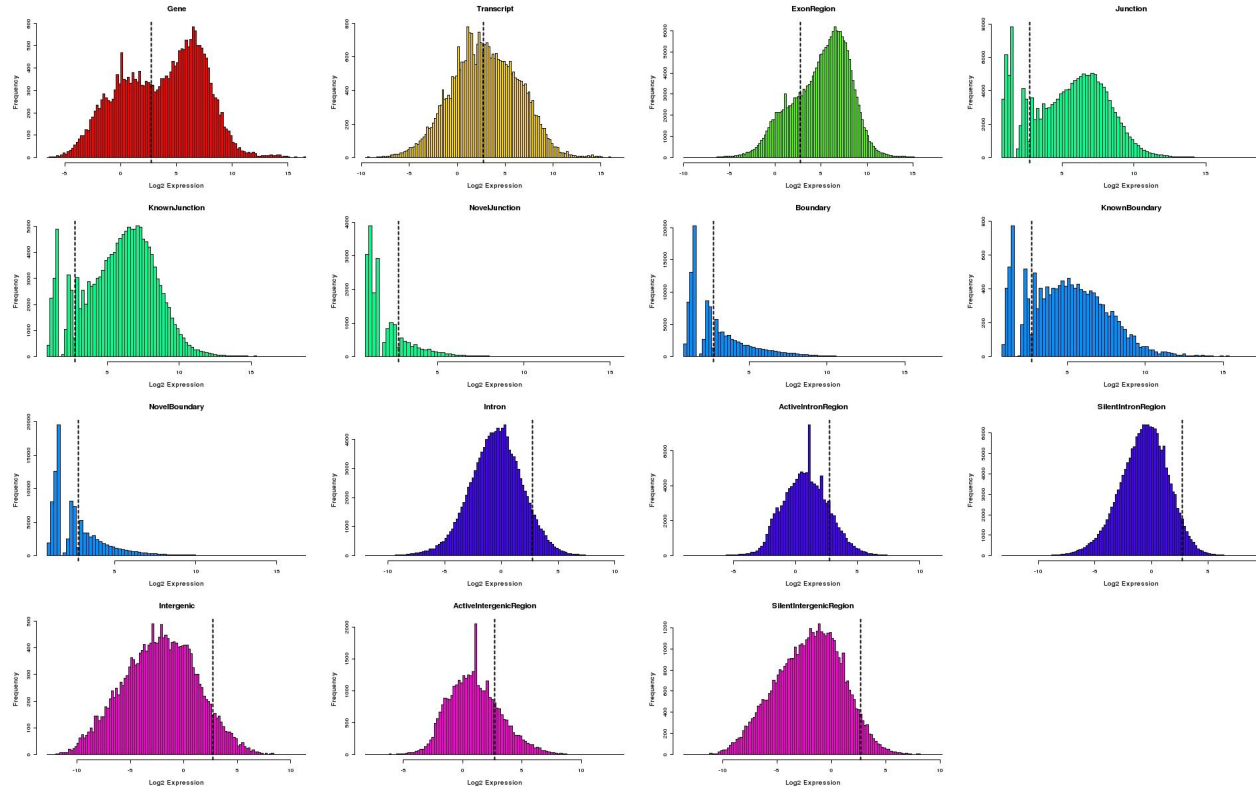
an·a·lyt·ics \,a-nə-'li-tiks\

“the method of logical analysis”

source: Merriam-Webster



Looking at Distributions As A Service



Experiment Time!

Moral of the Story

There are hidden truths everywhere if you spend the time to look for them.



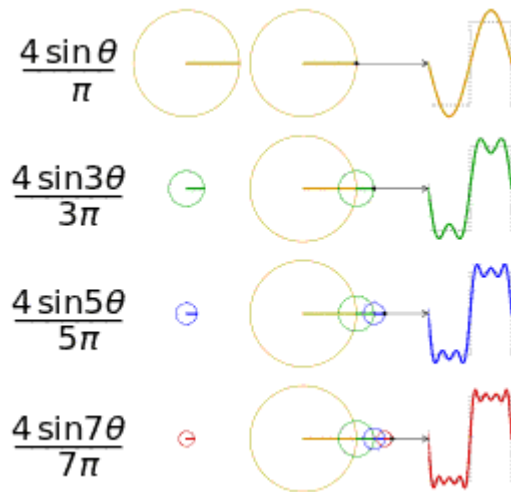
Facial Recognition



<http://www.theverge.com/2014/8/8/5982727/face-wreck-how-advanced-tech-comes-up-short-for-police>

Facial Recognition

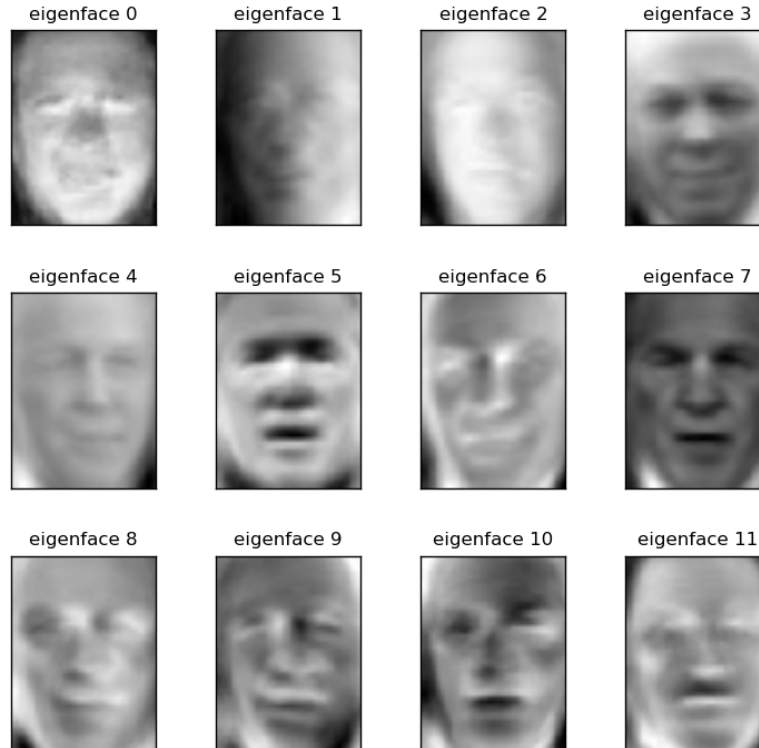
Fourier Series



Each arrow starts at the vertical sum of all the arrows on its left. And the purple dot is the sum of all six. The arrows represent the amplitudes of sine functions with different peak-values and frequencies. They are the first six terms of a Fourier series derived from the square wave motion of the blue dot, which transitions between only two amplitudes every 2 seconds.

source: http://en.wikipedia.org/wiki/Fourier_series

Eigenfaces



source: http://scikit-learn.org/0.10/auto_examples/applications/face_recognition.html

Eigenfaces

predicted: Powell
true: Powell



predicted: Sharon
true: Sharon



predicted: Bush
true: Bush



predicted: Bush
true: Bush



predicted: Bush
true: Blair



predicted: Rumsfeld
true: Rumsfeld



predicted: Rumsfeld
true: Rumsfeld



predicted: Rumsfeld
true: Blair



predicted: Bush
true: Bush



predicted: Powell
true: Powell



predicted: Chavez
true: Chavez

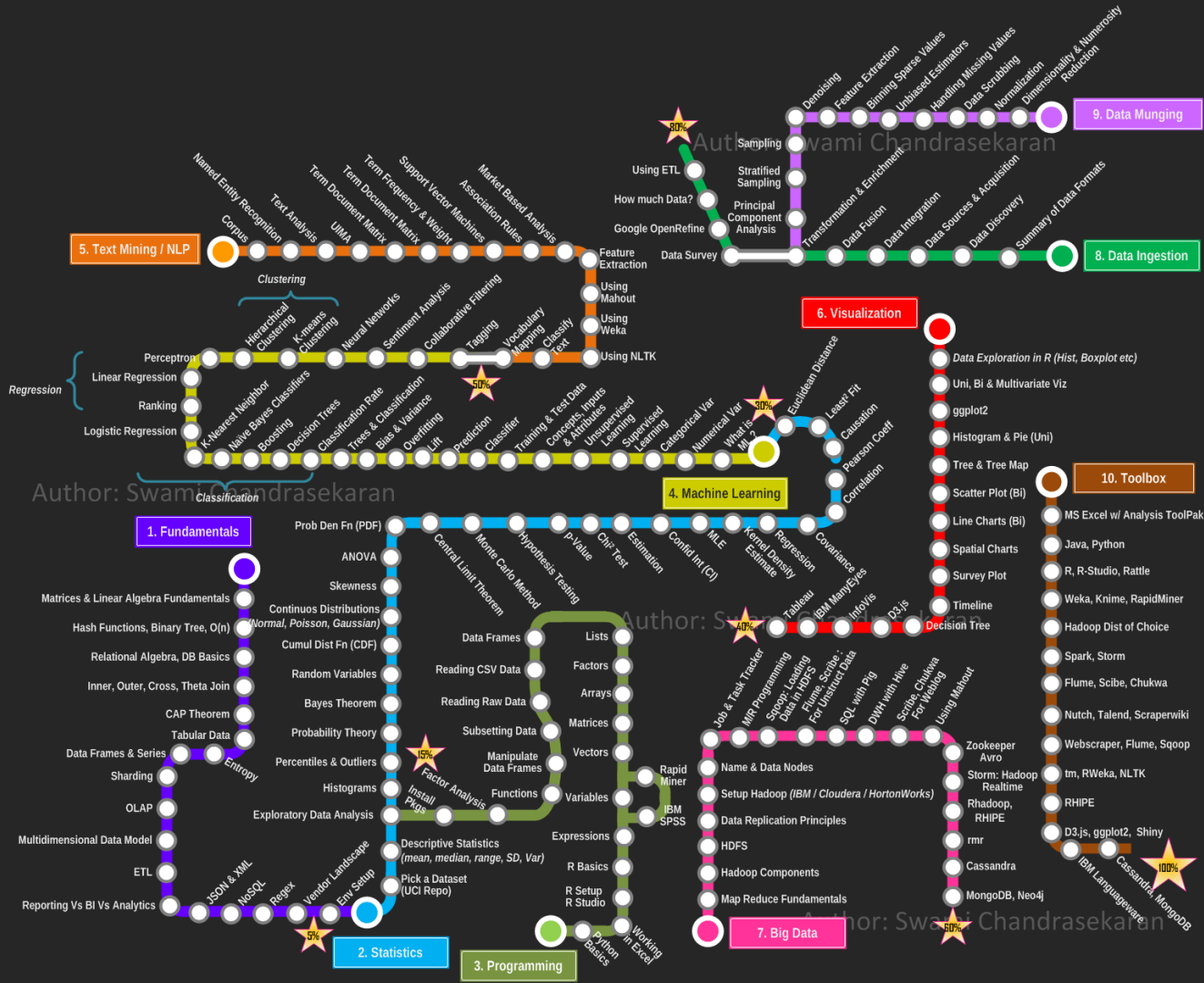


predicted: Rumsfeld
true: Powell



Data Science





In Context

$$15 - 8 =$$

How do you know when a student has learned this skill?



In Context

Get it right just once?

Was it a lucky guess?

Is it rote memorization?

Can she do “eight less than fifteen”?

Get it right every time?

Was she distracted?

Did she use regrouping?

$$15 - 8 =$$

Did she miss it because she was sick that day?

Can she teach someone else?

Will she be able to do it again in 2 weeks?

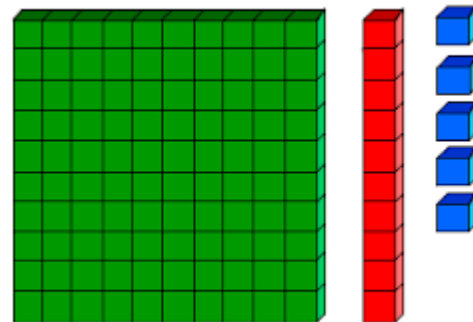
Does this mean she can subtract double digits?

How fast did she answer it?

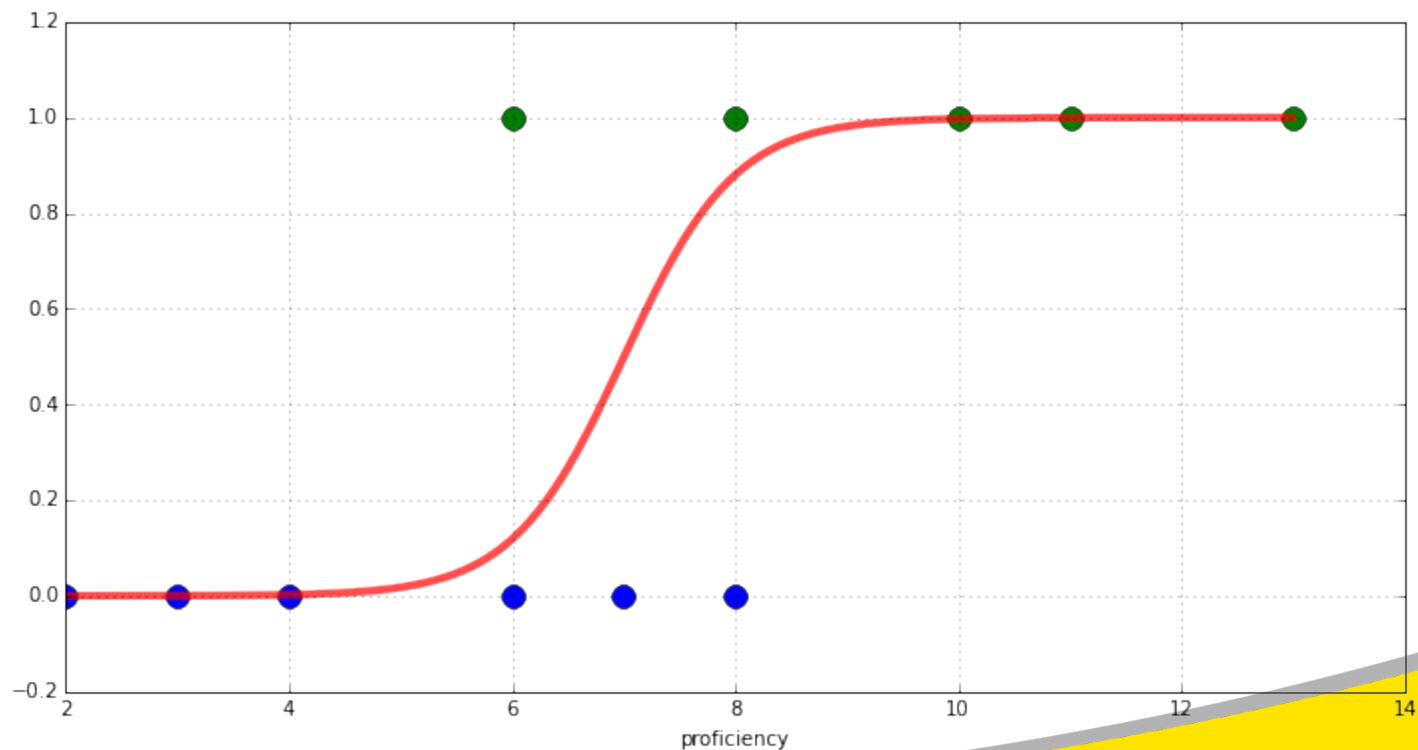
What about $15 - 9$? or $8 - 15$?

Was it in base 10 or base 9 or maybe base 13?

Can she do it with these?



In Context - Logistic Regression



“Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.”

- George Box



Random

- SAT/PSAT/GRE
- Measurement Noise
- Process Noise
- Partial Credit
- Bloom's Taxonomy
- Interoperability
- Socioeconomics
- Robust

