

# ML Week

Bonjour !

Jeff Abrahamson

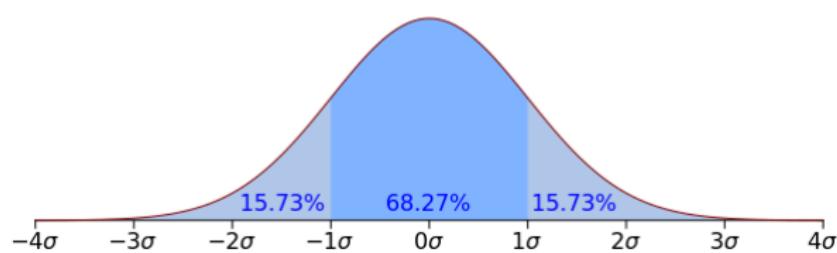
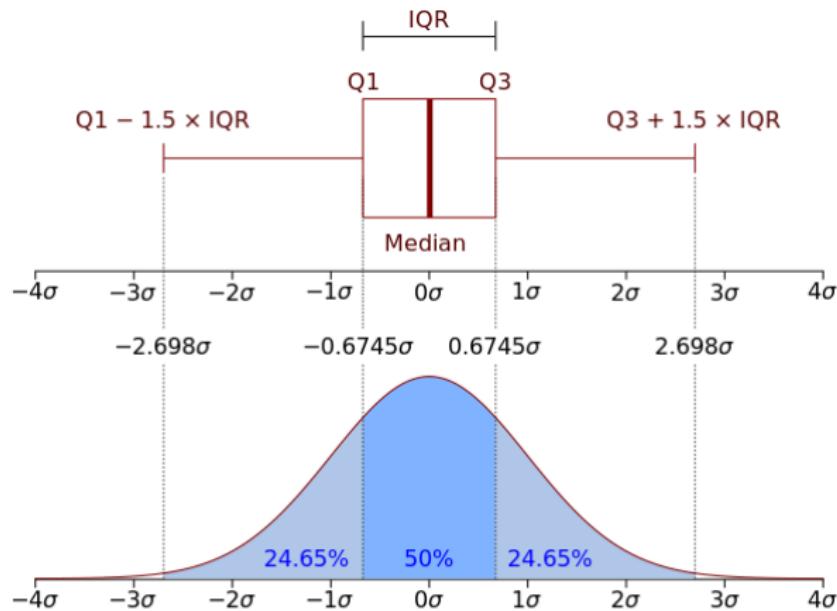
15 janvier 2016

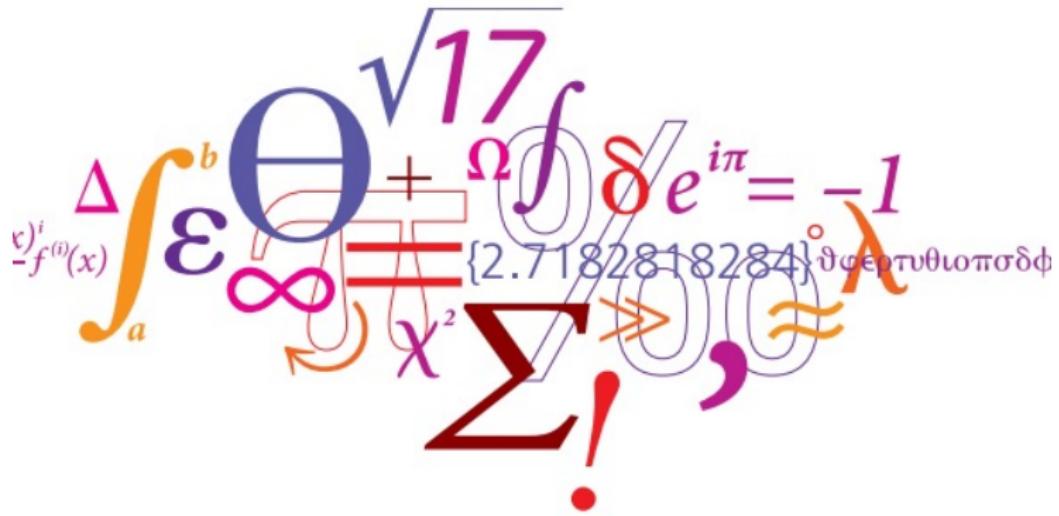
# **Machine Learning**

# Apprentissage automatique

# **Data Science**

# **Statistiques**

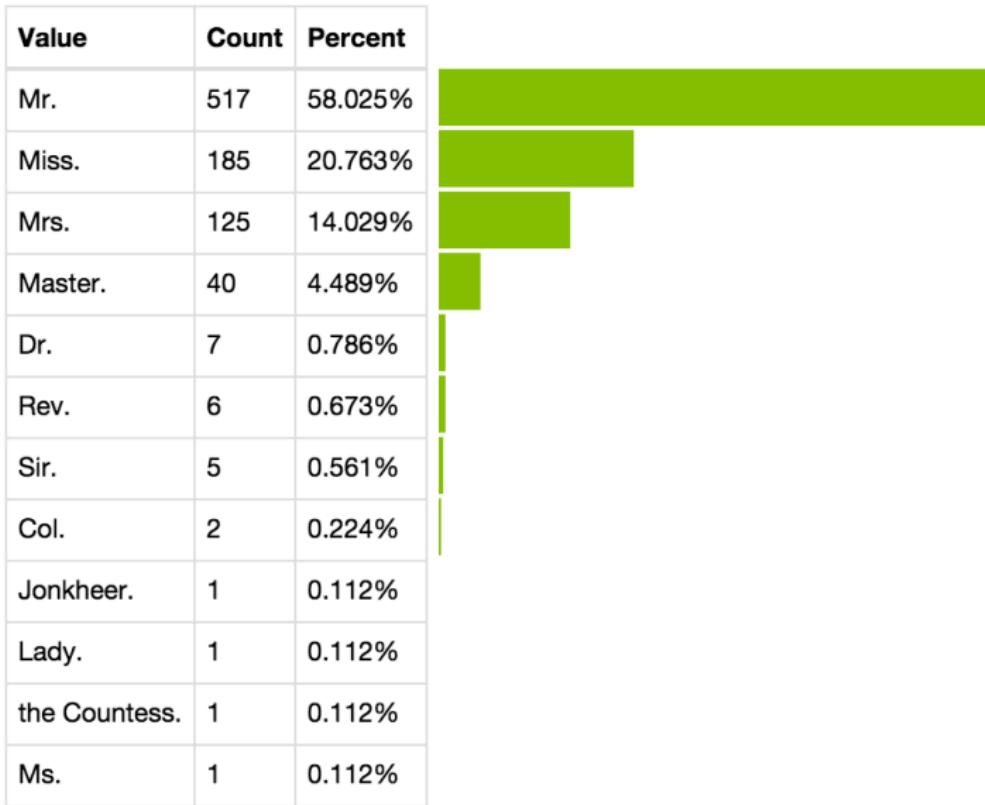




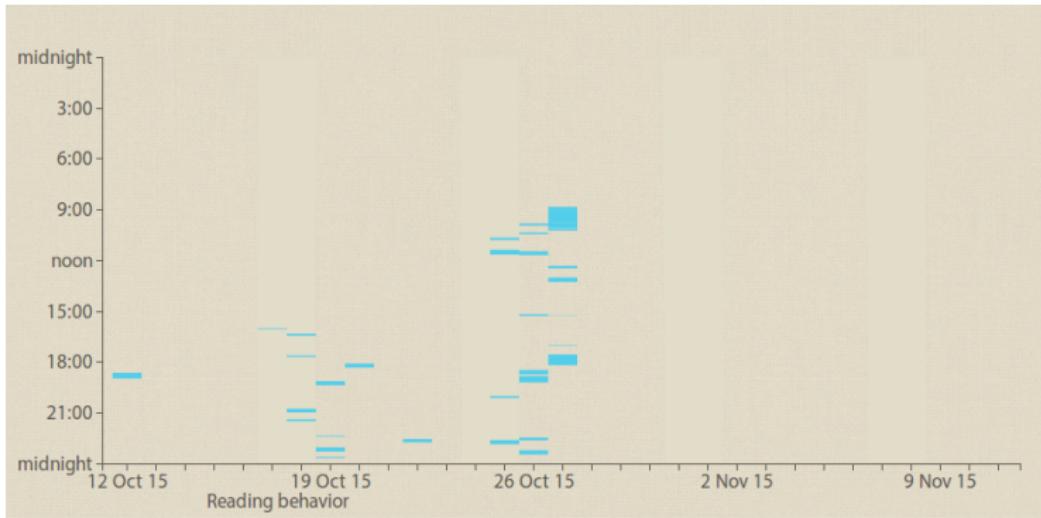
**Espace vectoriel**

**Critères**

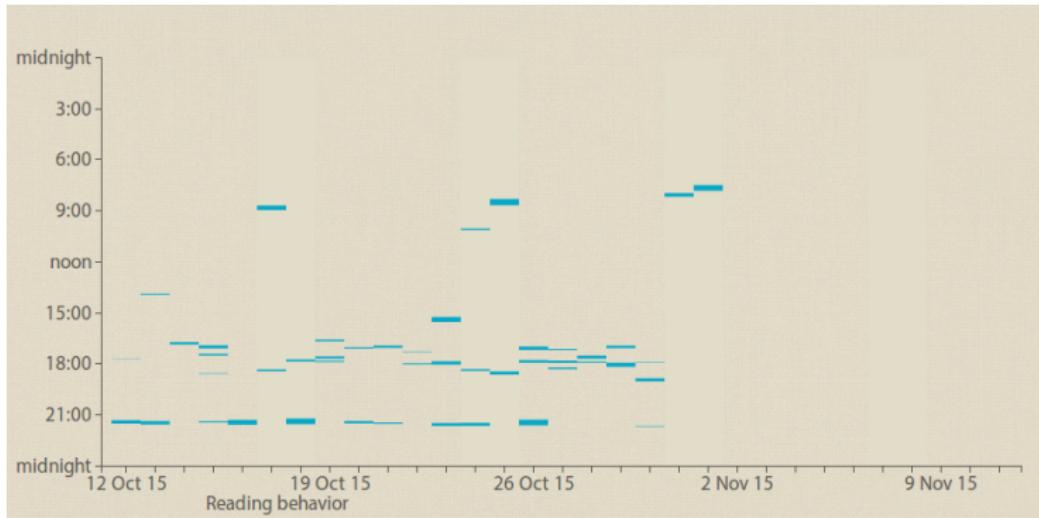
**Feature engineering**



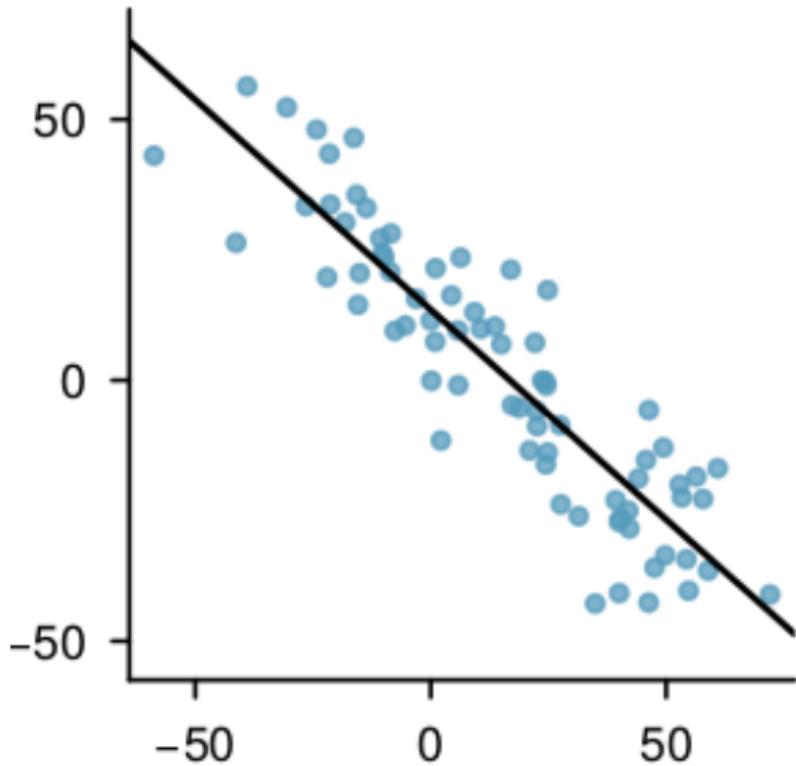
Kaggle

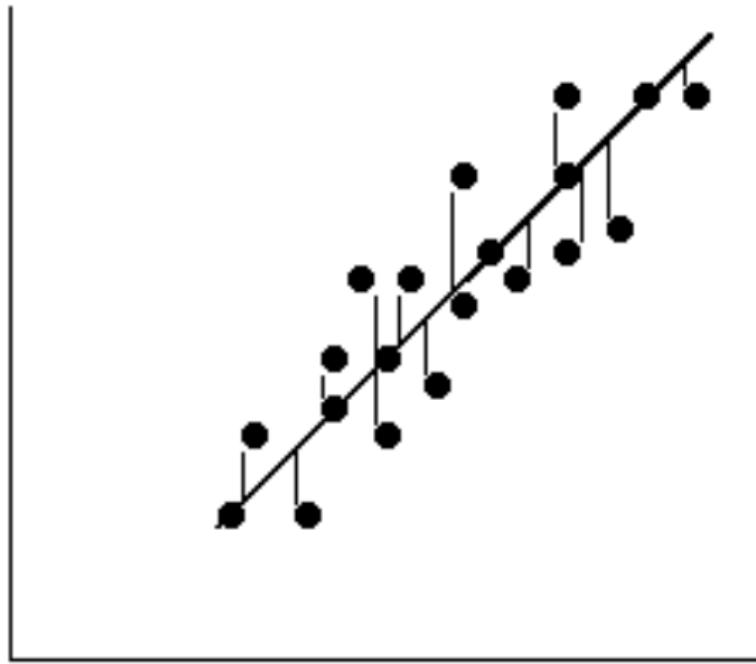


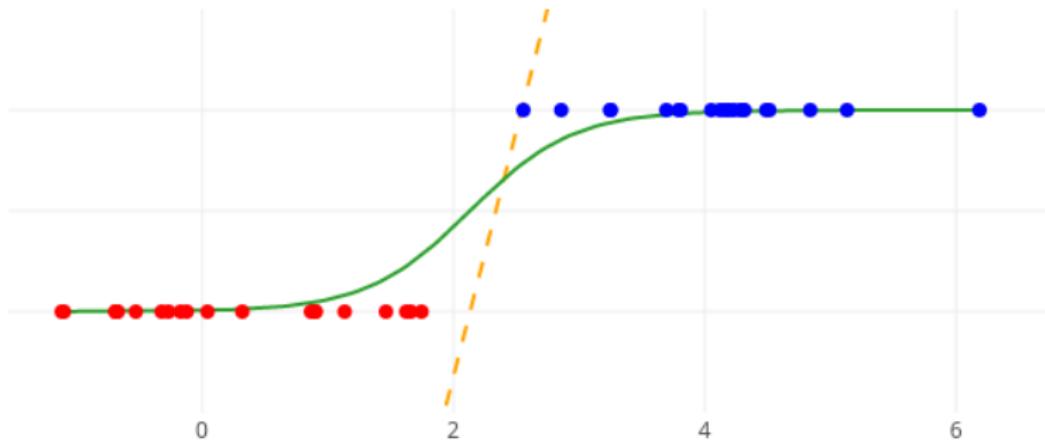
*Jellybooks*

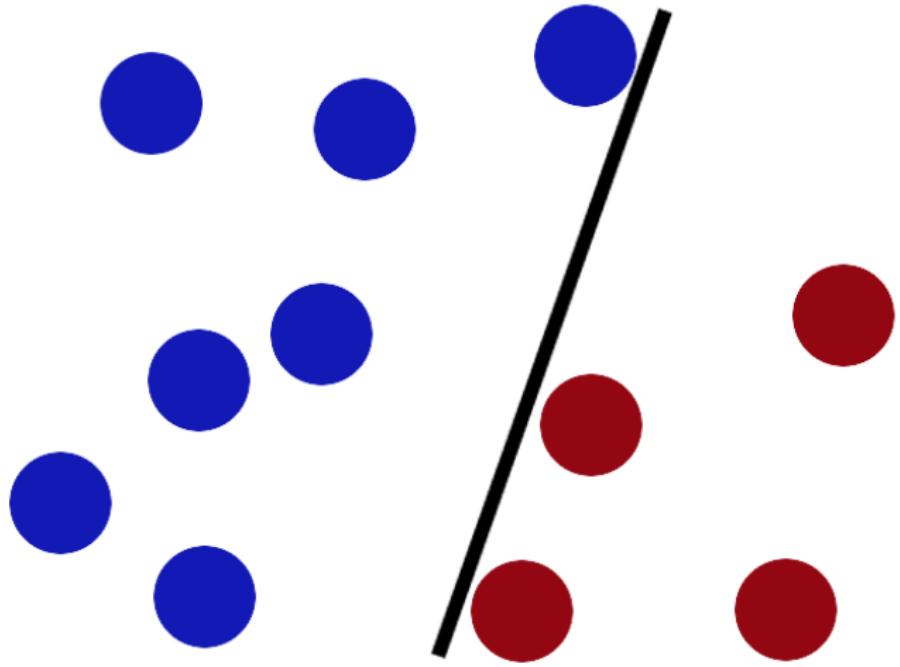


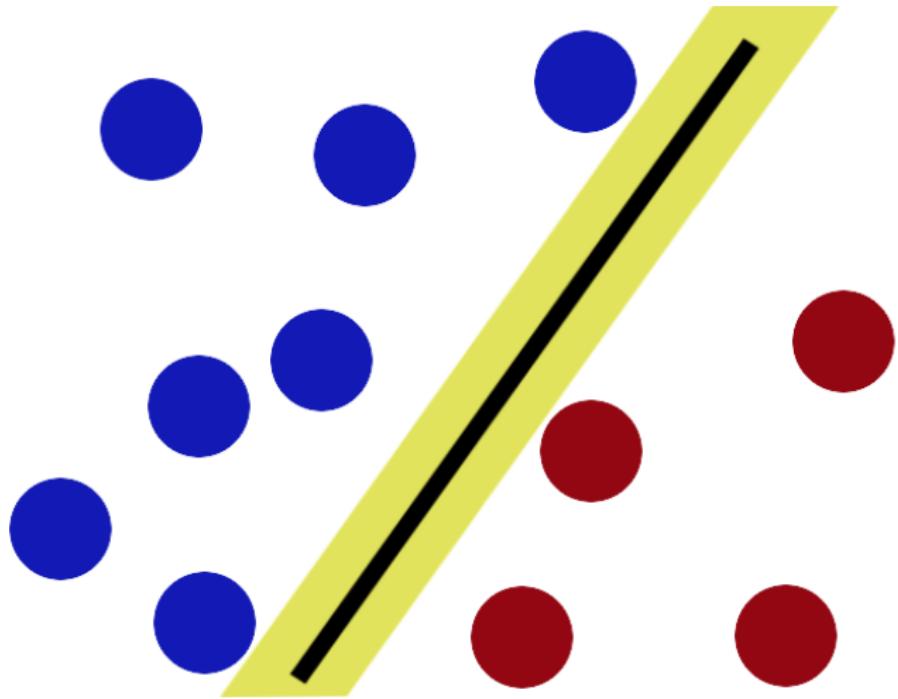
Jellybooks

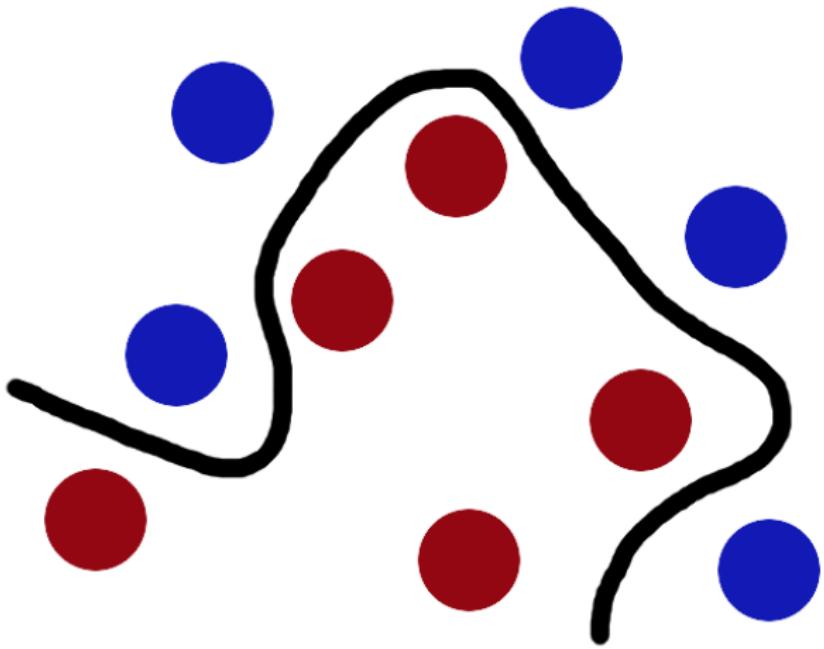


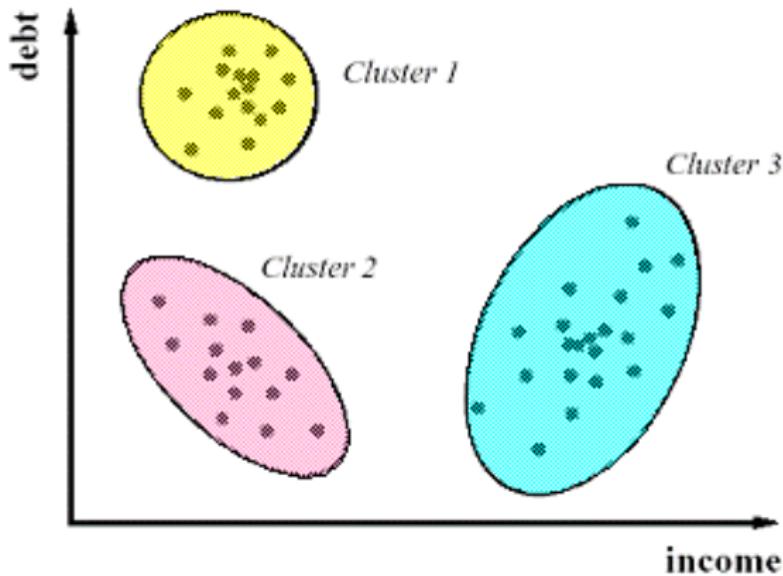


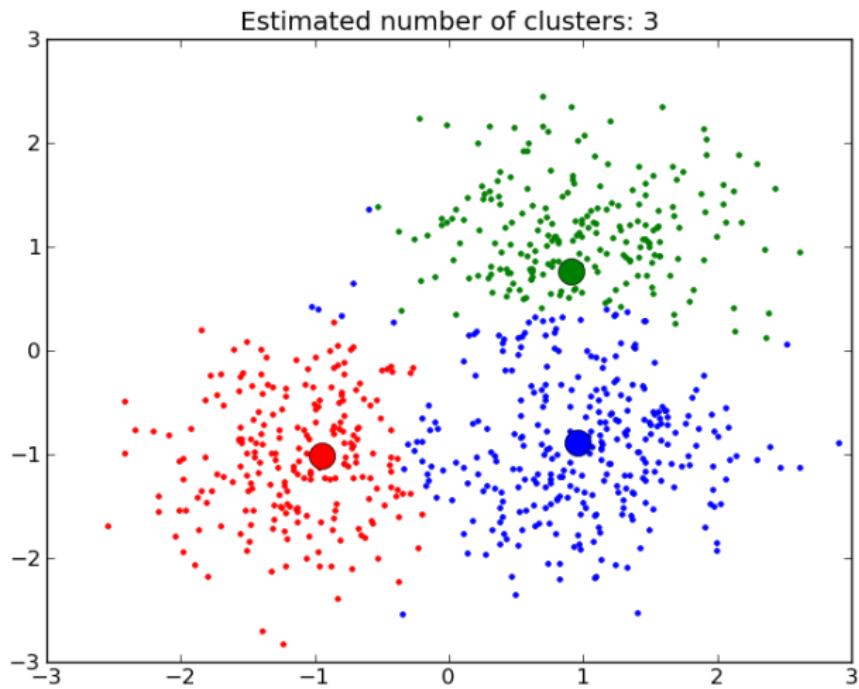


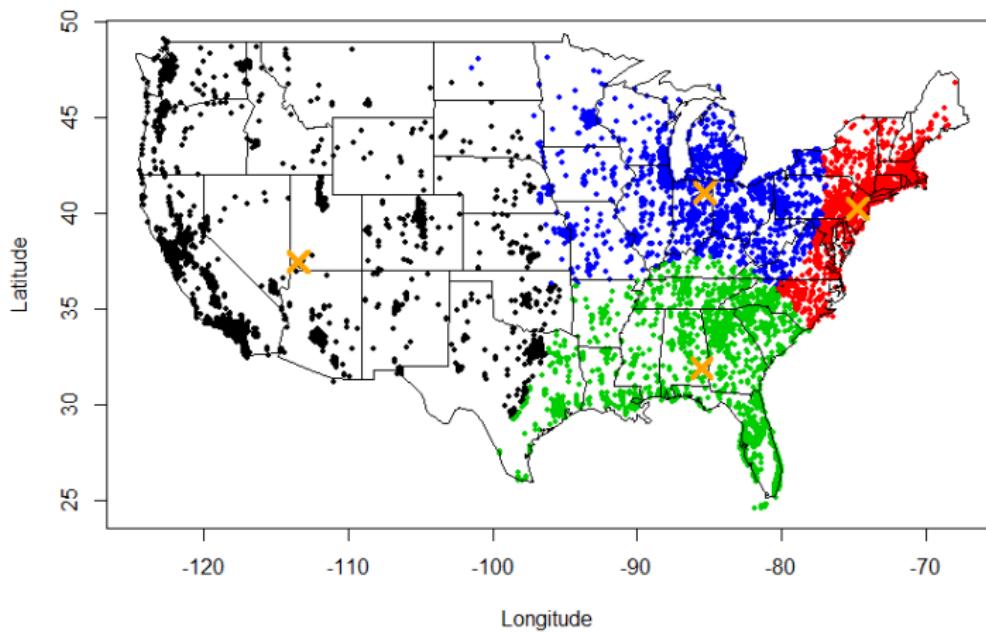




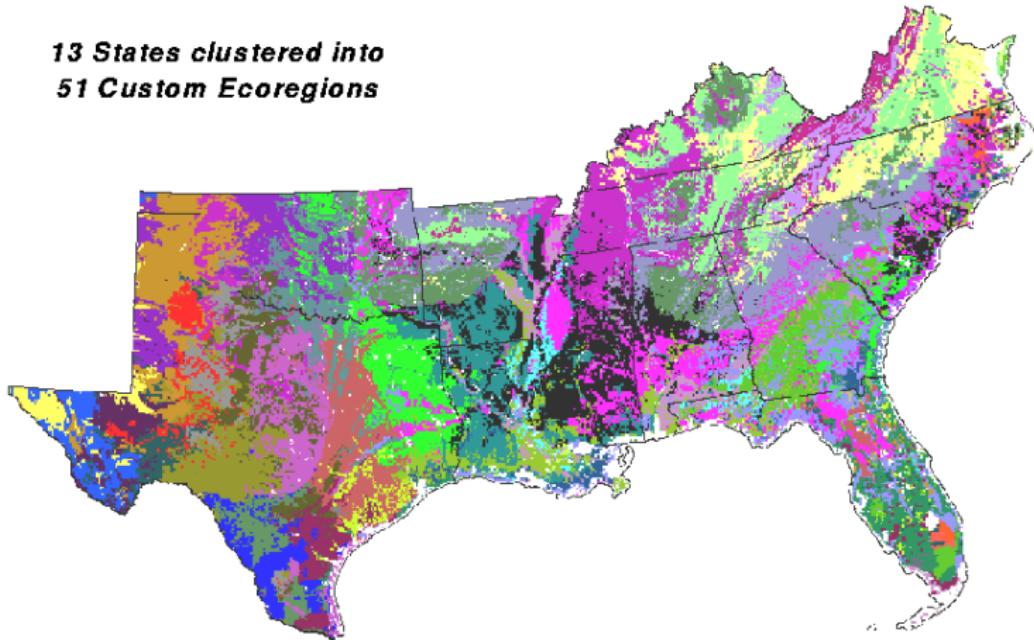








***13 States clustered into  
51 Custom Ecoregions***





*Eddie Bell @ Lyst*



*Eddie Bell @ Lyst*



*Eddie Bell @ Lyst*



*Eddie Bell @ Lyst*

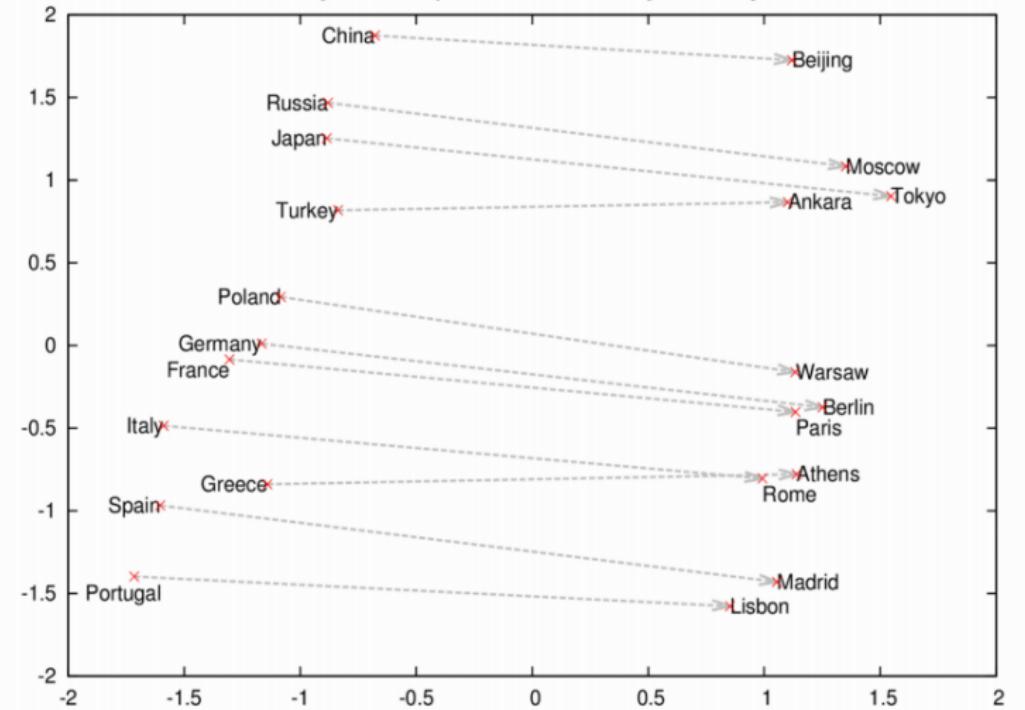


-1	0	+1
-2	0	+2
-1	0	+1

-1	-2	-1
0	0	0
+1	+2	+1



### Country and Capital Vectors Projected by PCA



Term	Similarity	
	"shift"	0.933104
	"gown"	0.887743
	"skirt"	0.881672
	"bandage"	0.880162
	"midi"	0.869786

**Similar to 'dress'**

*Eddie Bell @ Lyst*



a group of young girls standing next  
to each other on the beach



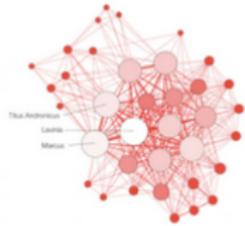
A clock tower with a clock on top of it

*Google?*



## A bunch of bananas hanging from a tree

*Google?*



**TITUS ANDRONICUS**  
Number of characters **36** | 50% Network density



**ROMEO AND JULIET**  
Number of characters **41** | 37% Network density



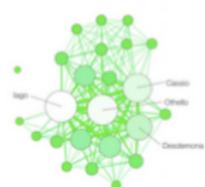
**JULIUS CAESAR**  
Number of characters **46** | 34% Network density



**HAMLET**  
Number of characters **37** | 39% Network density

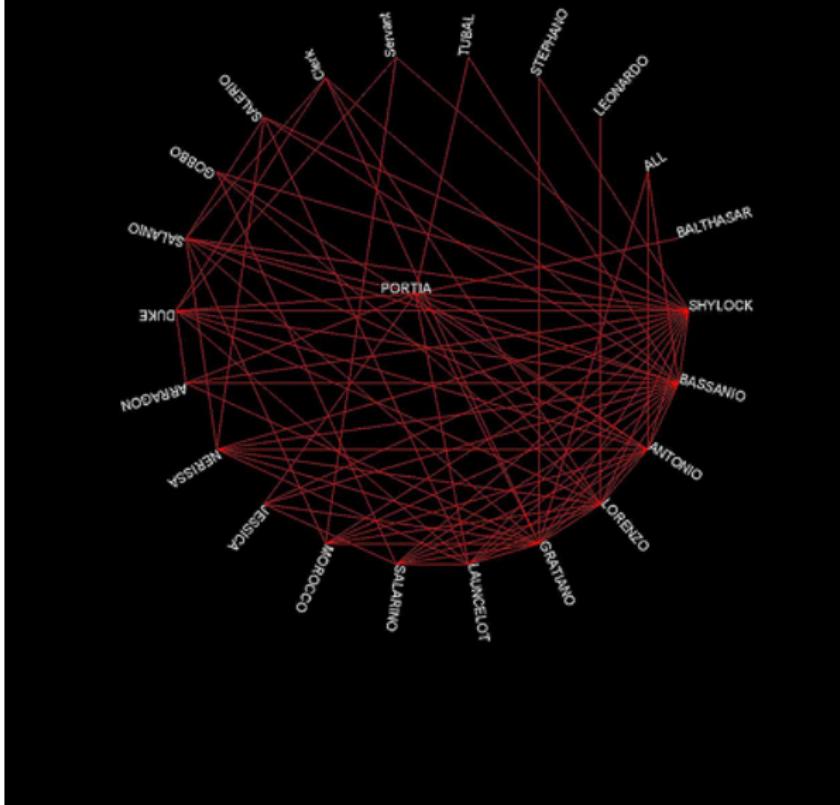


**TROILUS AND CRESSIDA**  
Number of characters **35** | 40% Network density



**OTHELLO**  
Number of characters **24** | 55% Network density

[https://flowingdata.com/2015/12/30/  
shakespeare-tragedies-as-network-graphs/](https://flowingdata.com/2015/12/30/shakespeare-tragedies-as-network-graphs/)



## Act 4, Scene 1

**Perseus** And I am nothing else to black Fox than  
**Perseus** Crimson is the colour, I like it not.

**Perseus** And therefore have I little talked of love.  
**Perseus** Look, see, how crimson the lady towards my self,  
**Perseus** Needs me not let us go.  
**Perseus** That she do, we muste be aske  
**Perseus** What she sayes, I thinke shee shal say  
**Perseus** What shee sayes.  
**Perseus** Come to me, make confession to her father,  
**Perseus** By a secret that I alwaies keepe  
**Perseus** By no means to her, for shee can  
**Perseus** I will shew her your countrey  
**Perseus** So what? I will see her face  
**Perseus** Being spidey before your fayre, than to your fayre  
**Perseus** How comly shee is, and how comly shee  
**Perseus** The traynes haue gonne my selfe by that  
**Perseus** This dayes a morninges sunne, that raynes  
**Perseus** And what i speake, i speake it to thy fayre  
**Perseus** My brewe come, and brewe assent to it  
**Perseus** Or shall I come to you at evening mass.  
**Perseus** My brewe serues me, peruse daugther, now  
**Perseus** Juliet, on Thursday early will I noise ye,  
**Perseus** just hope, just cure, just help.

**Perseus** I hear thou must, and nothing may prologue it.

Tell me not, friar, that thou  
hear'st of this.

Thou hast the strength of will to slay thyself.

Act 5, Scene 1

My dreams presage some joyful news at hand.

**Author:** Her body sleeps in Capels' monument.

**ANSWER** I will hence to-night.

*[uttering a low groan]*

Author: I do beseech you, sir, have patience.

Sustaining the Environment

get thee gone.

BRUNNEN

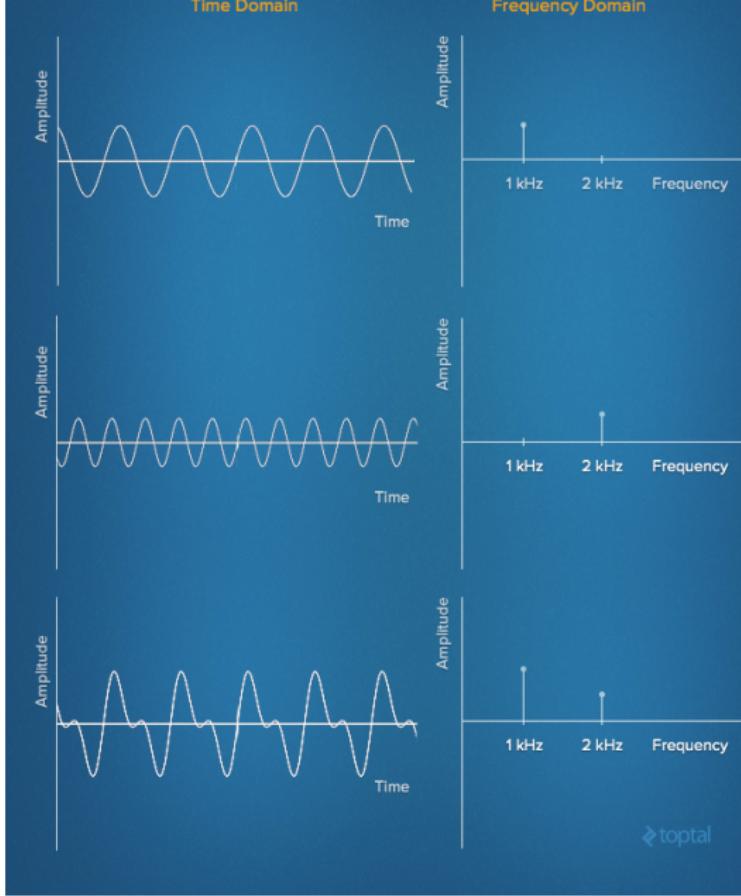
That the life-weary taker may fall dead.

is my theory. But Martha's law

famine is in thy cheeks.

### III Hypothesis

There is thy gold, worse poison to men's souls.



toptal

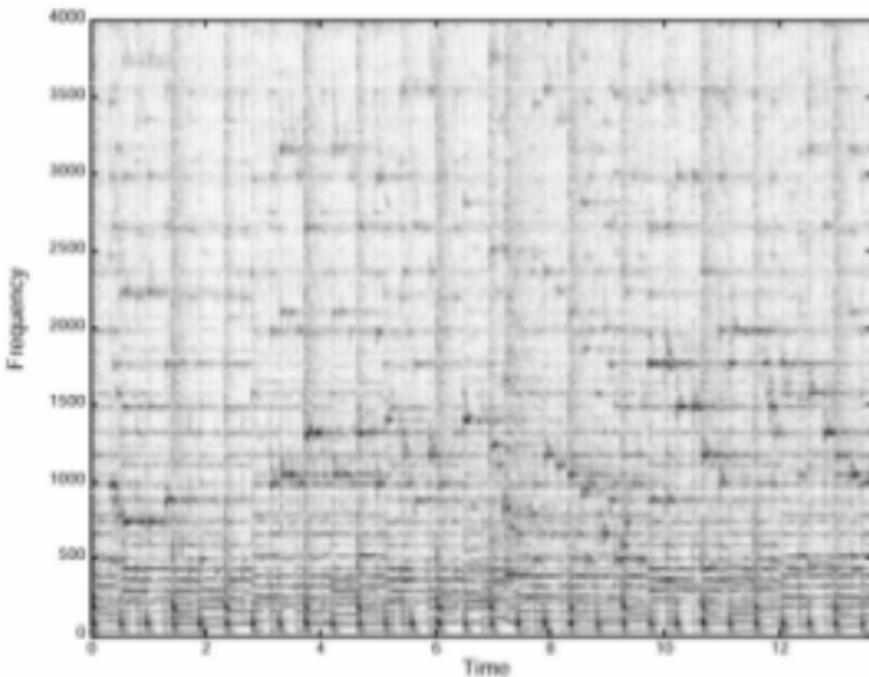


Fig. 1A - Spectrogram

*https:*

//www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf

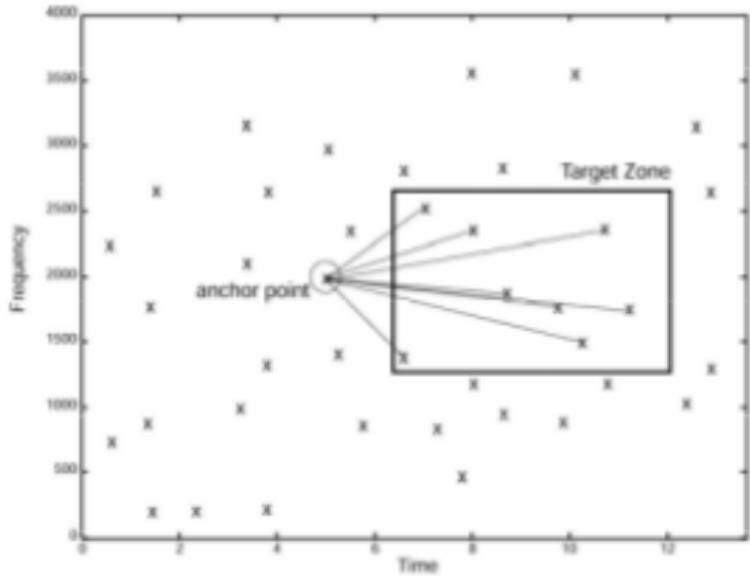
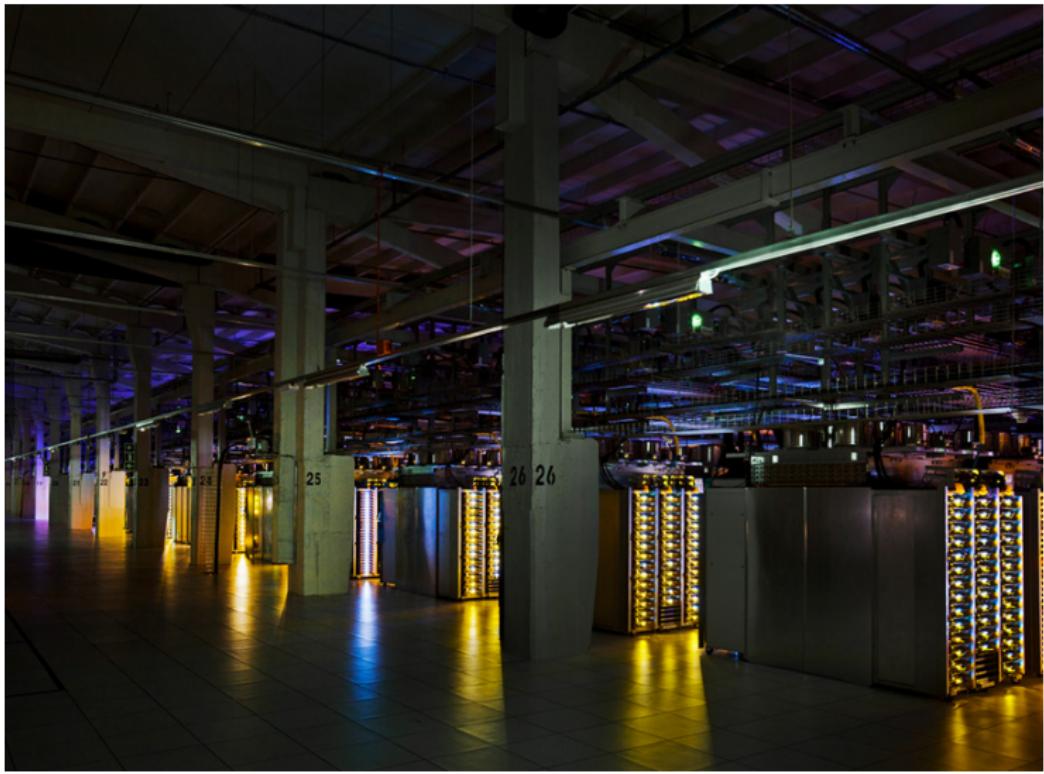


Fig. 1C - Combinatorial Hash Generation

[https:](https://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf)

//www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf



*Google*

## **Big data?**

*http://aadrake.com/*

*command-line-tools-can-be-235x-faster-than-your-hadoop-cluster.html*

# Un peu de maths...

- 100 critères
- int (*4 octets par critère*)
- 16 GB RAM
- $O(n \log n)$  en mémoire

$$400N + 400N \log(400N) < \frac{16 \cdot 10^9}{400}$$

# Un peu de maths...

- 100 critères
- int (*4 octets par critère*)
- 16 GB RAM
- $O(n \log n)$  en mémoire

$$400N + 400N \log(400N) < \frac{16 \cdot 10^9}{400}$$

$$N \approx 2 \cdot 10^6$$

# **Questions?**