

ML Week

Bonjour !

Jeff Abrahamson

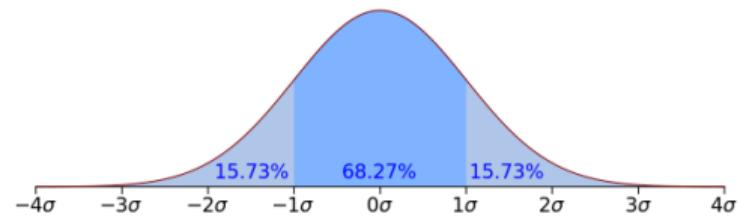
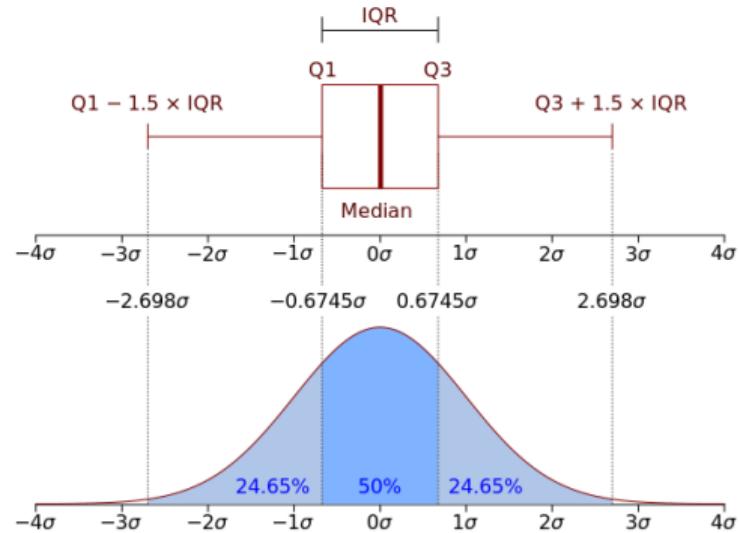
23–24 novembre 2016

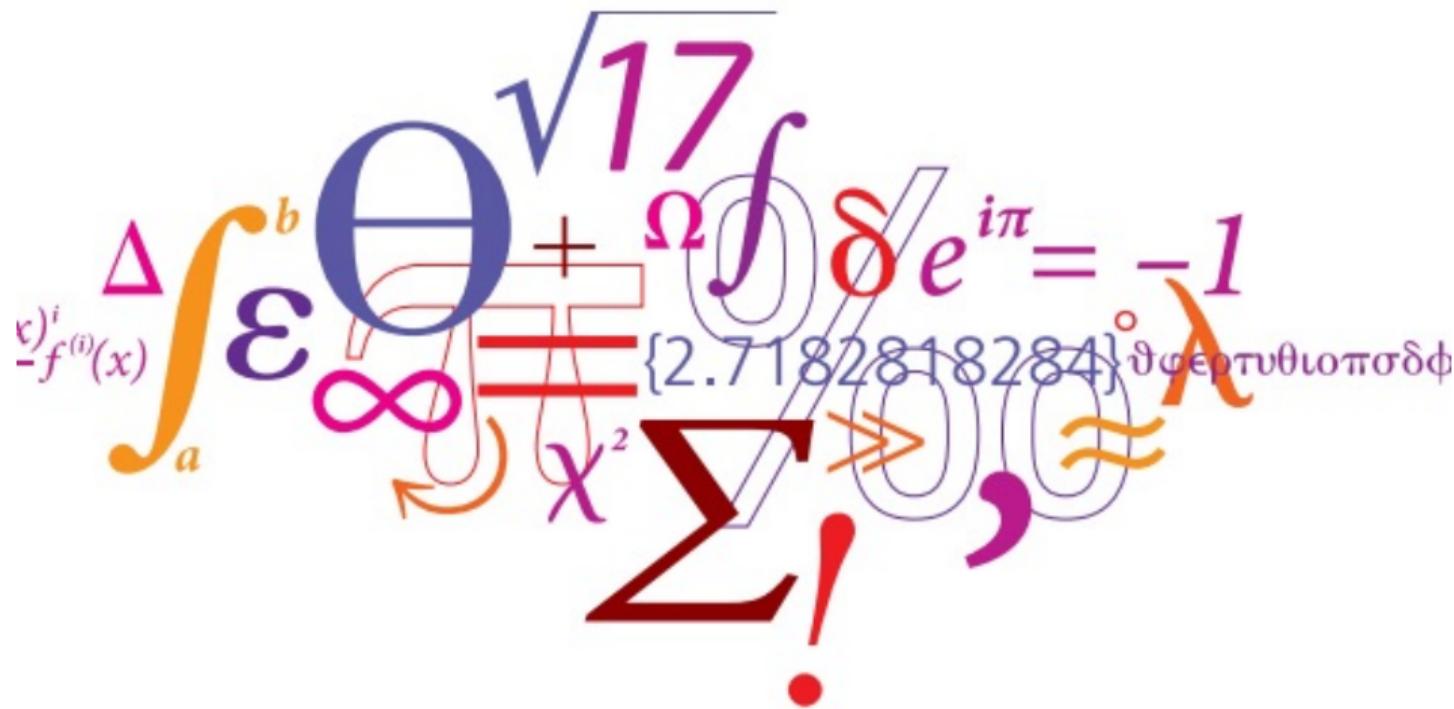
Machine Learning

Apprentissage automatique

Data Science

Statistiques

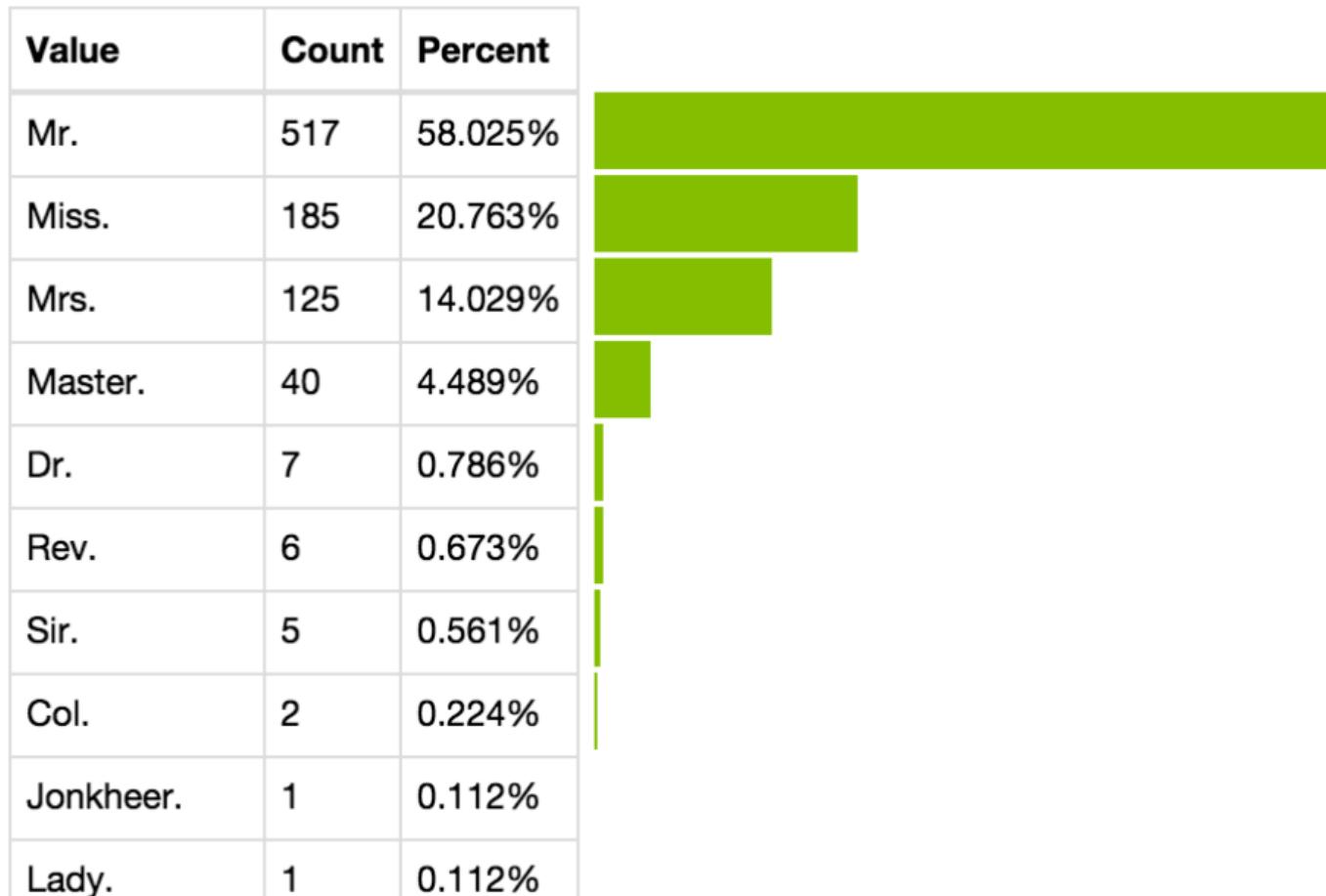


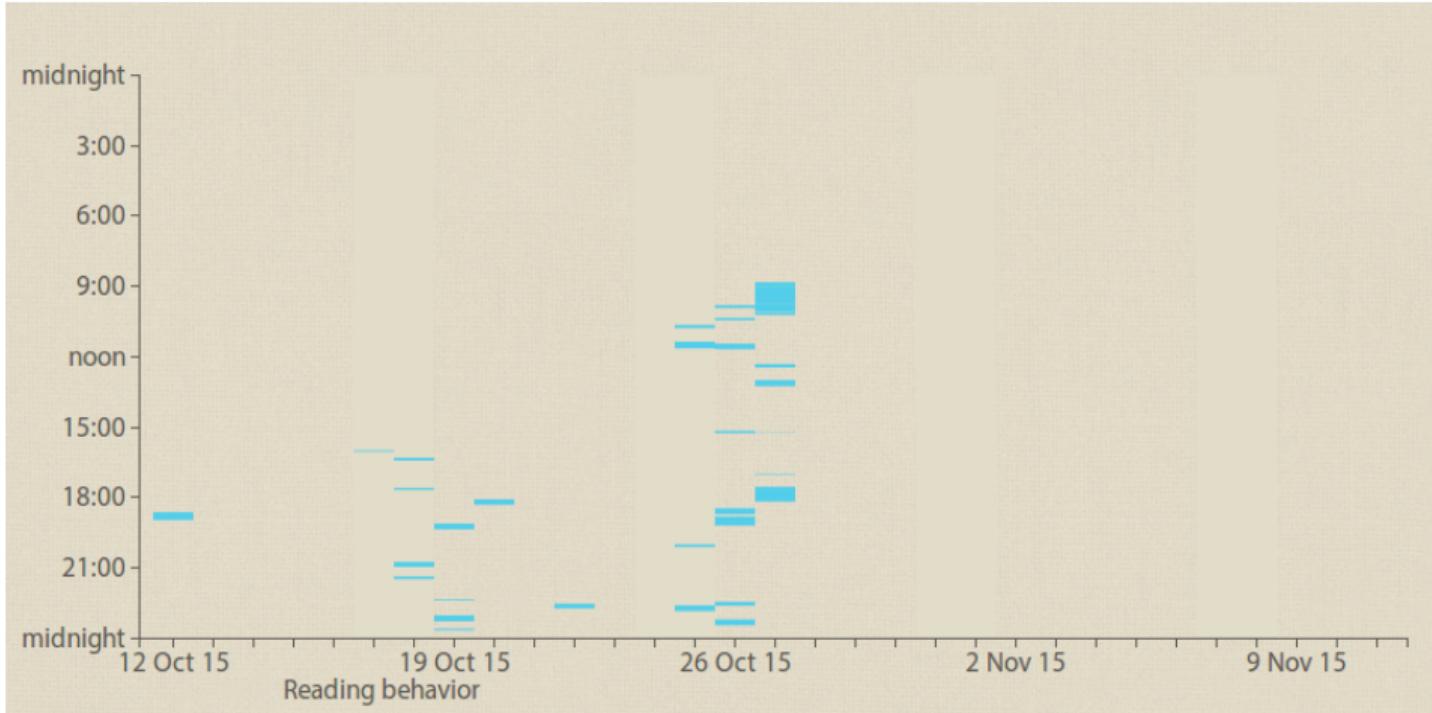


Espace vectoriel

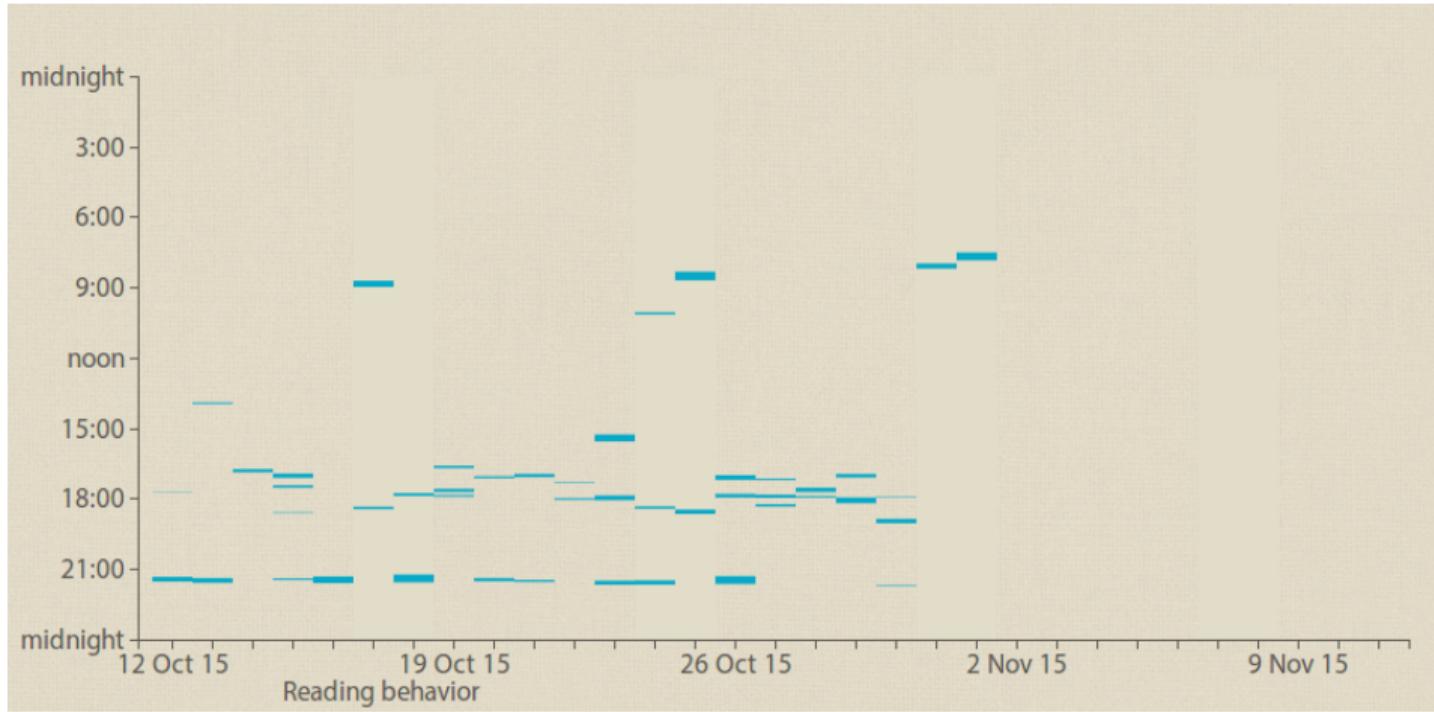
Critères

Feature engineering

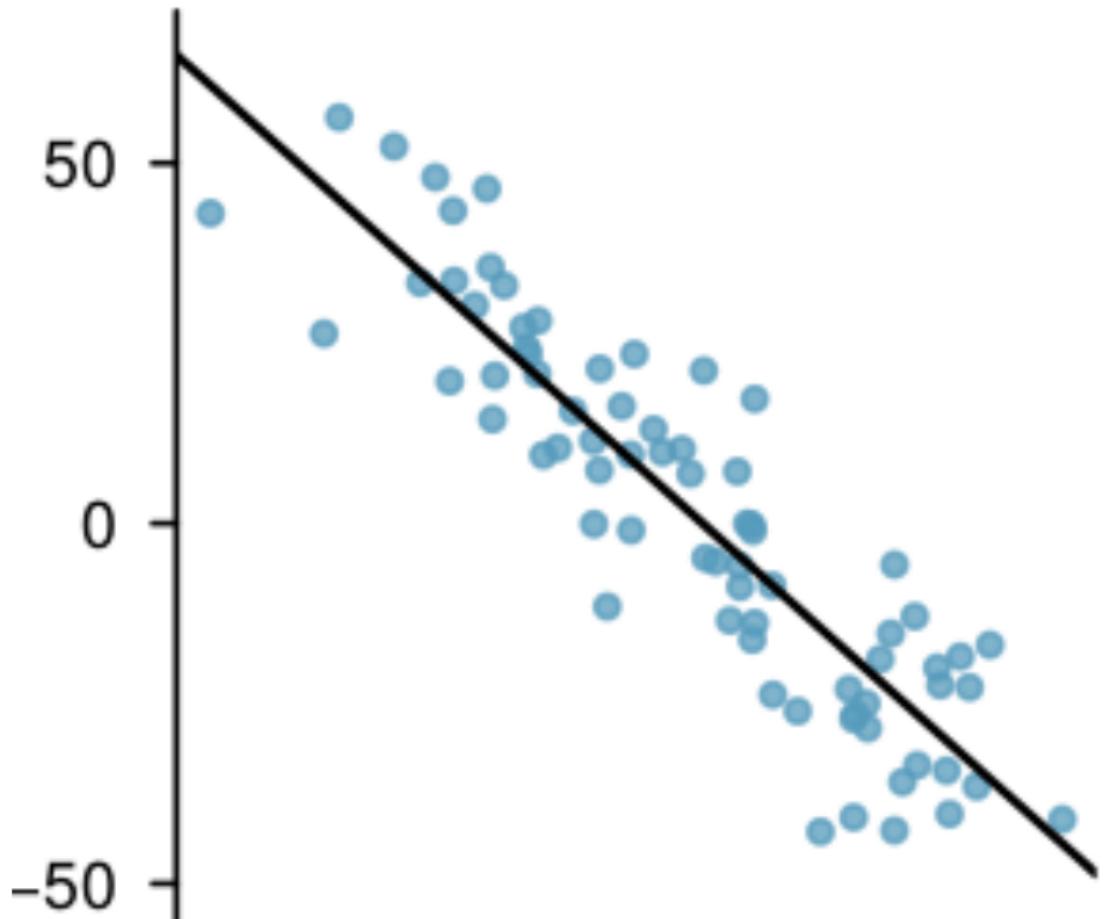


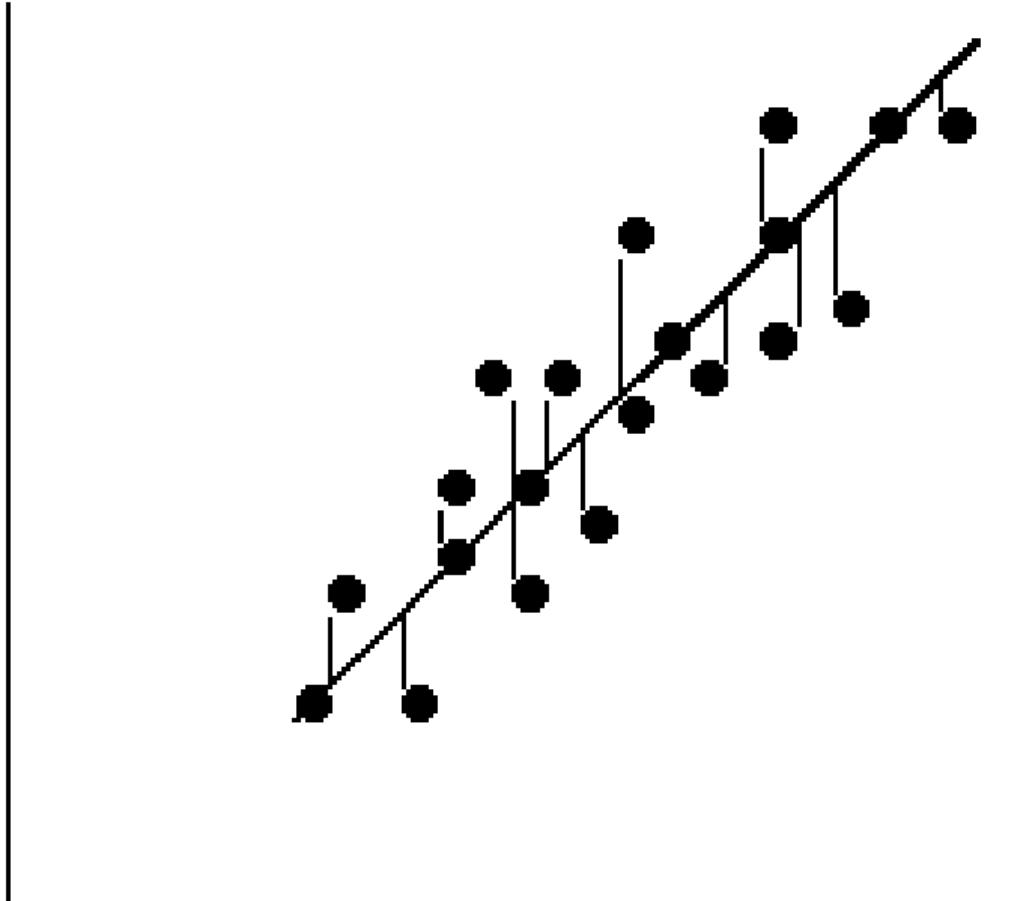


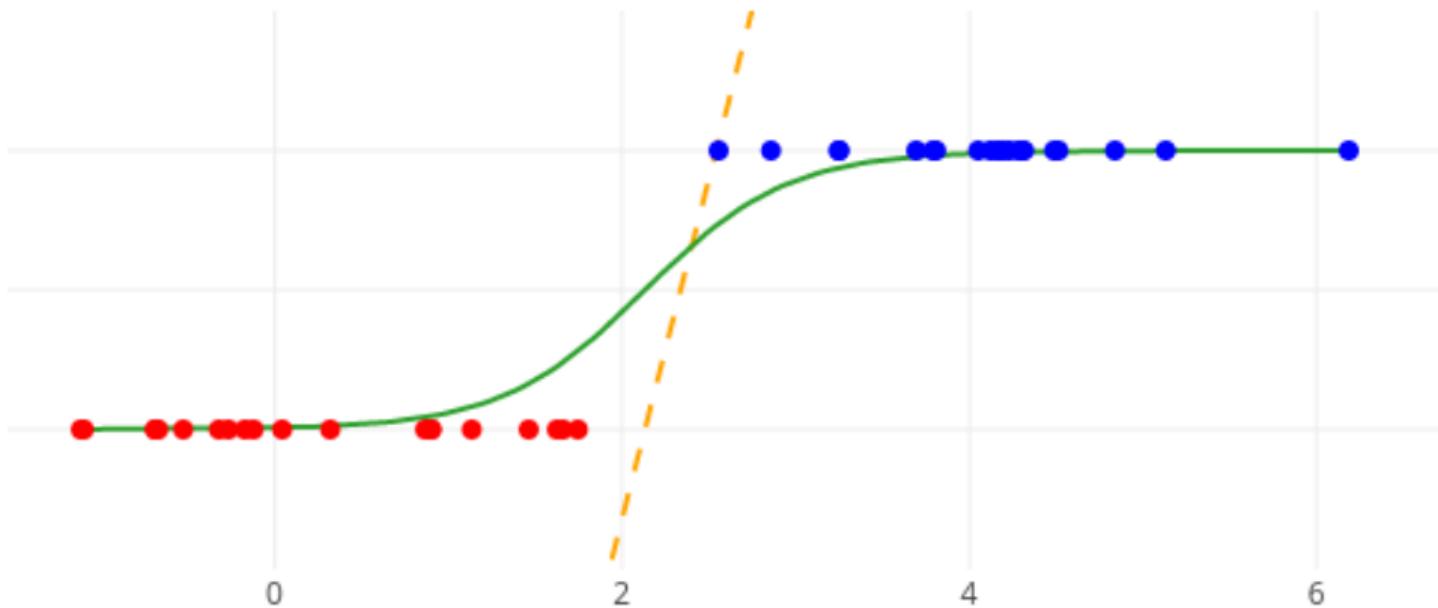
Jellybooks

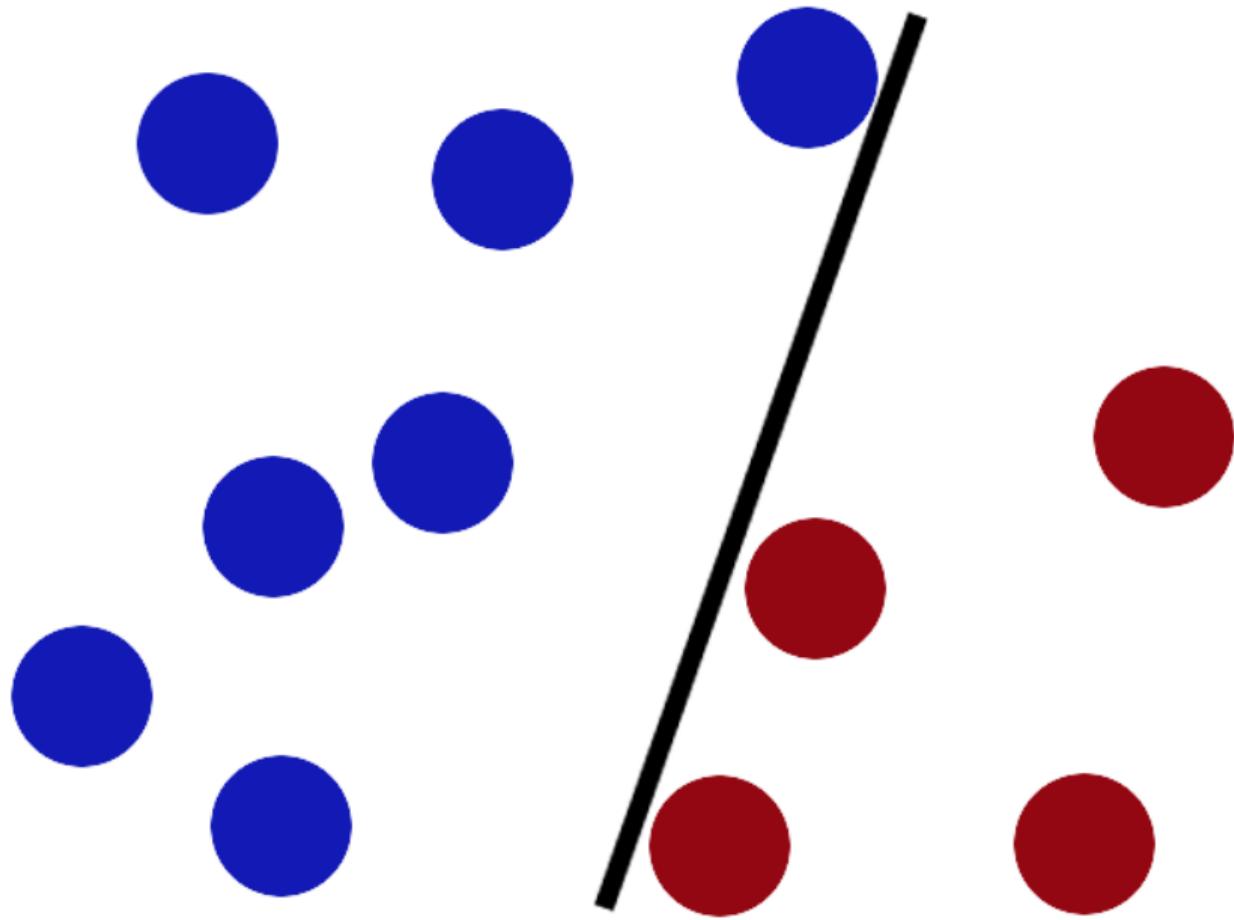


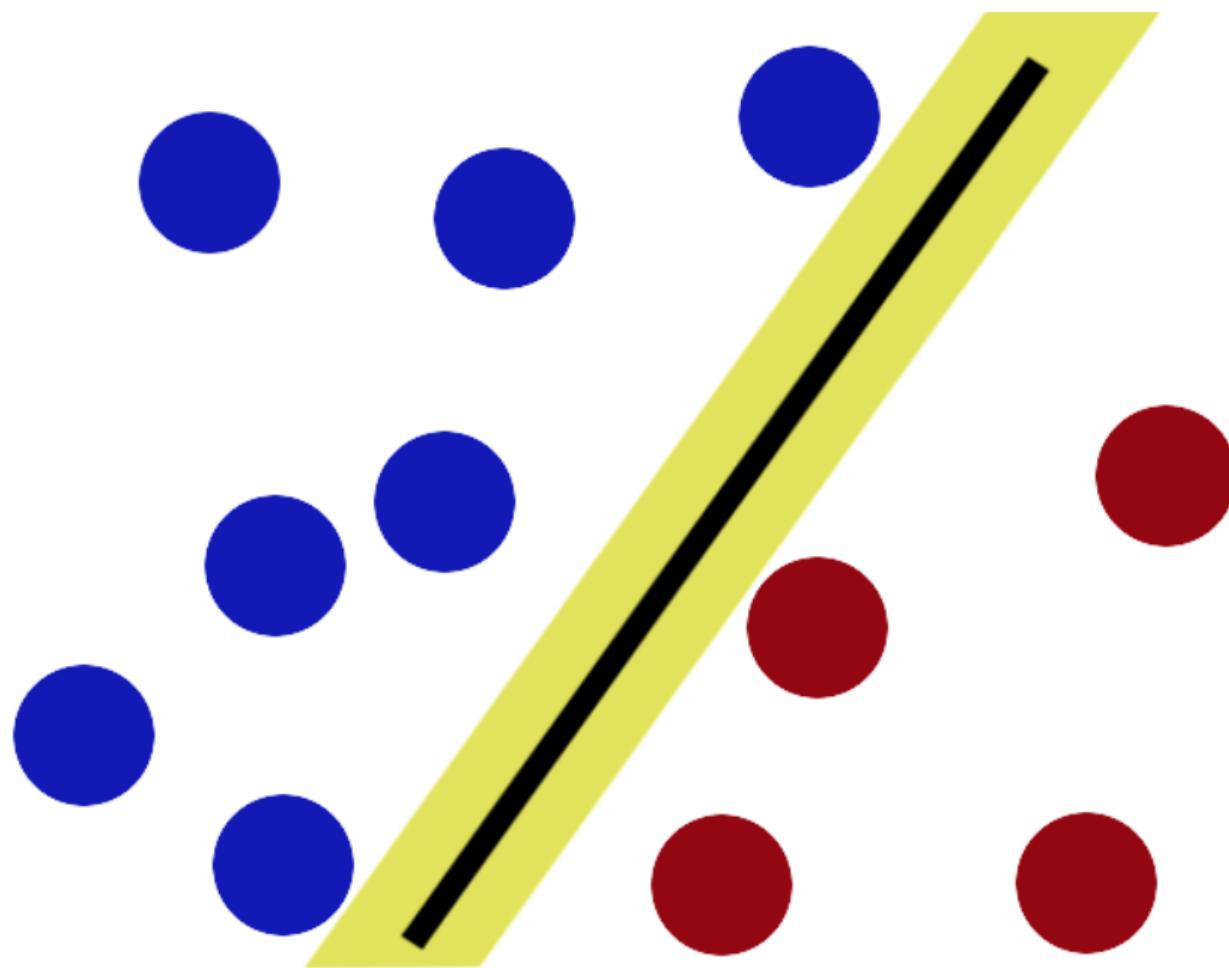
Jellybooks

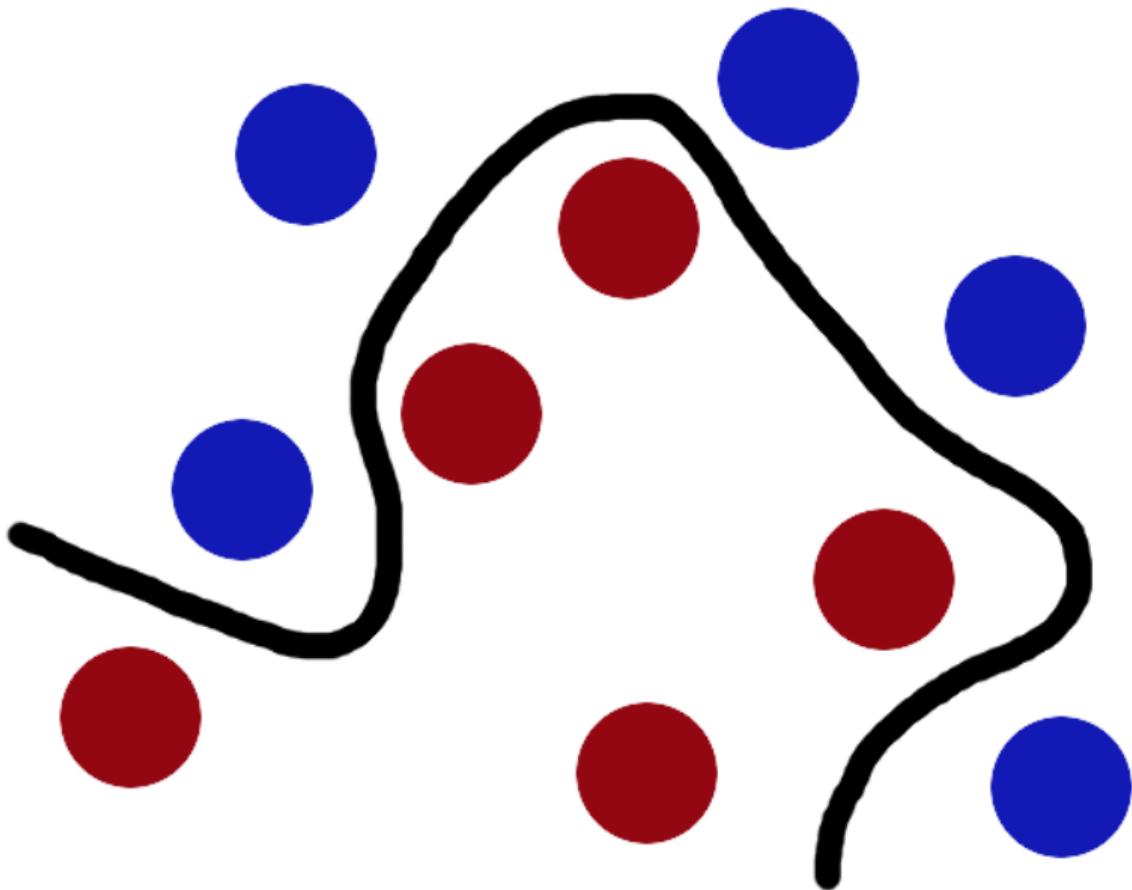


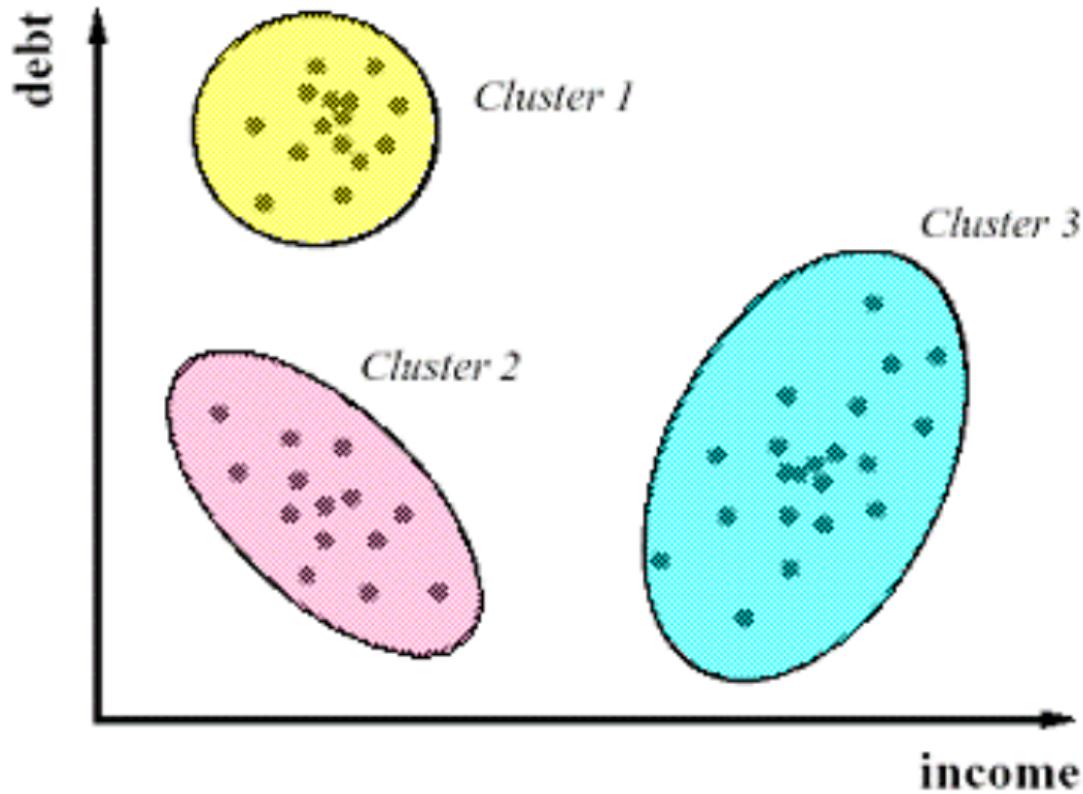




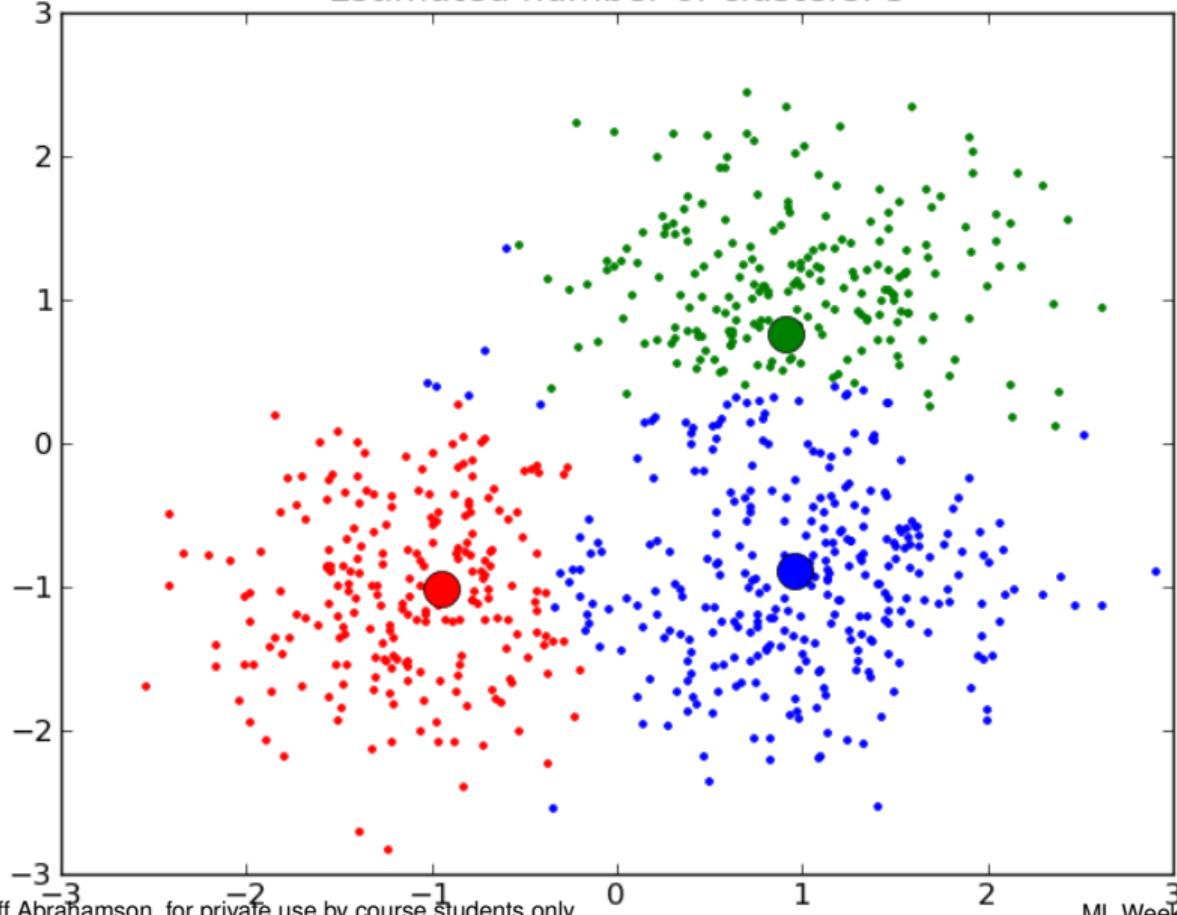


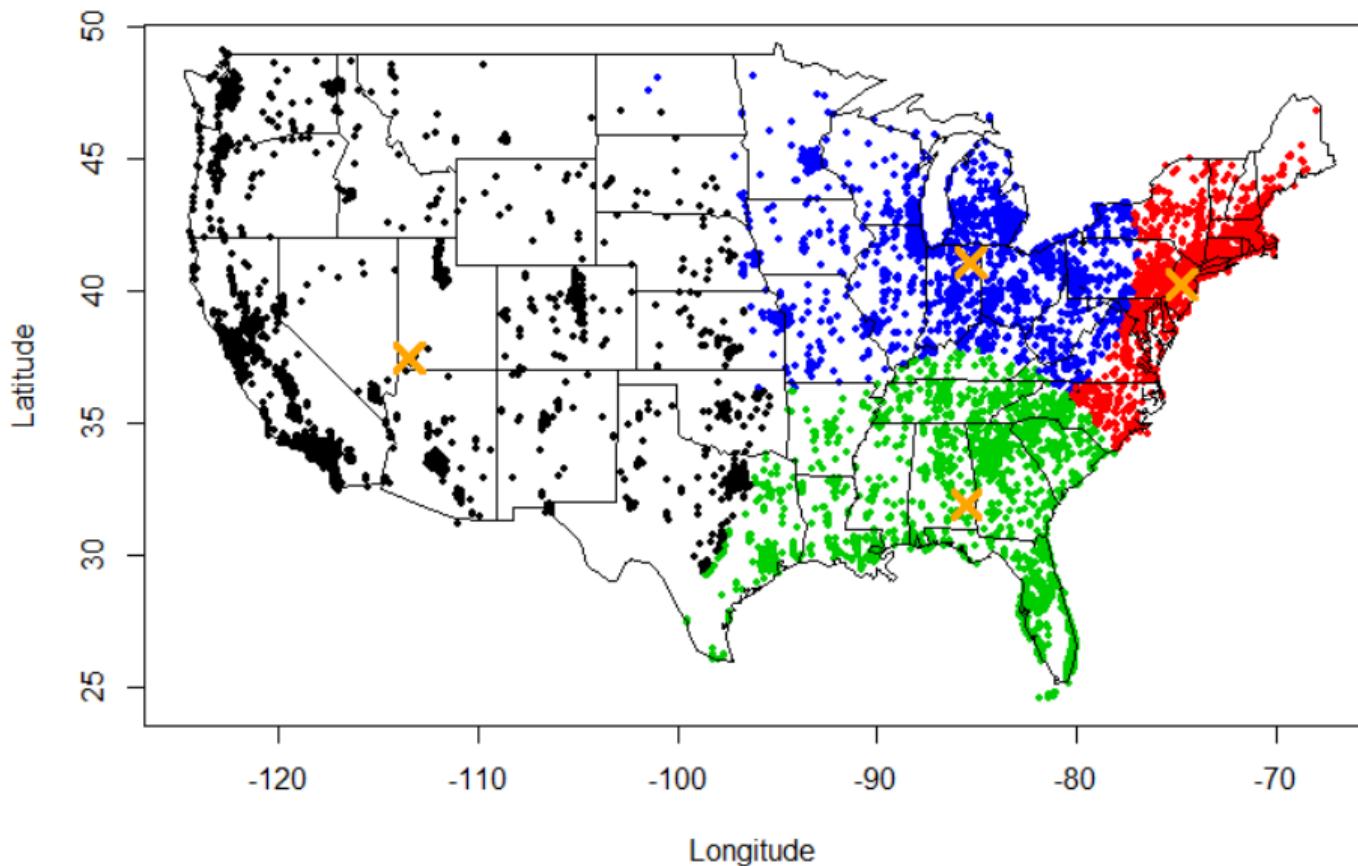




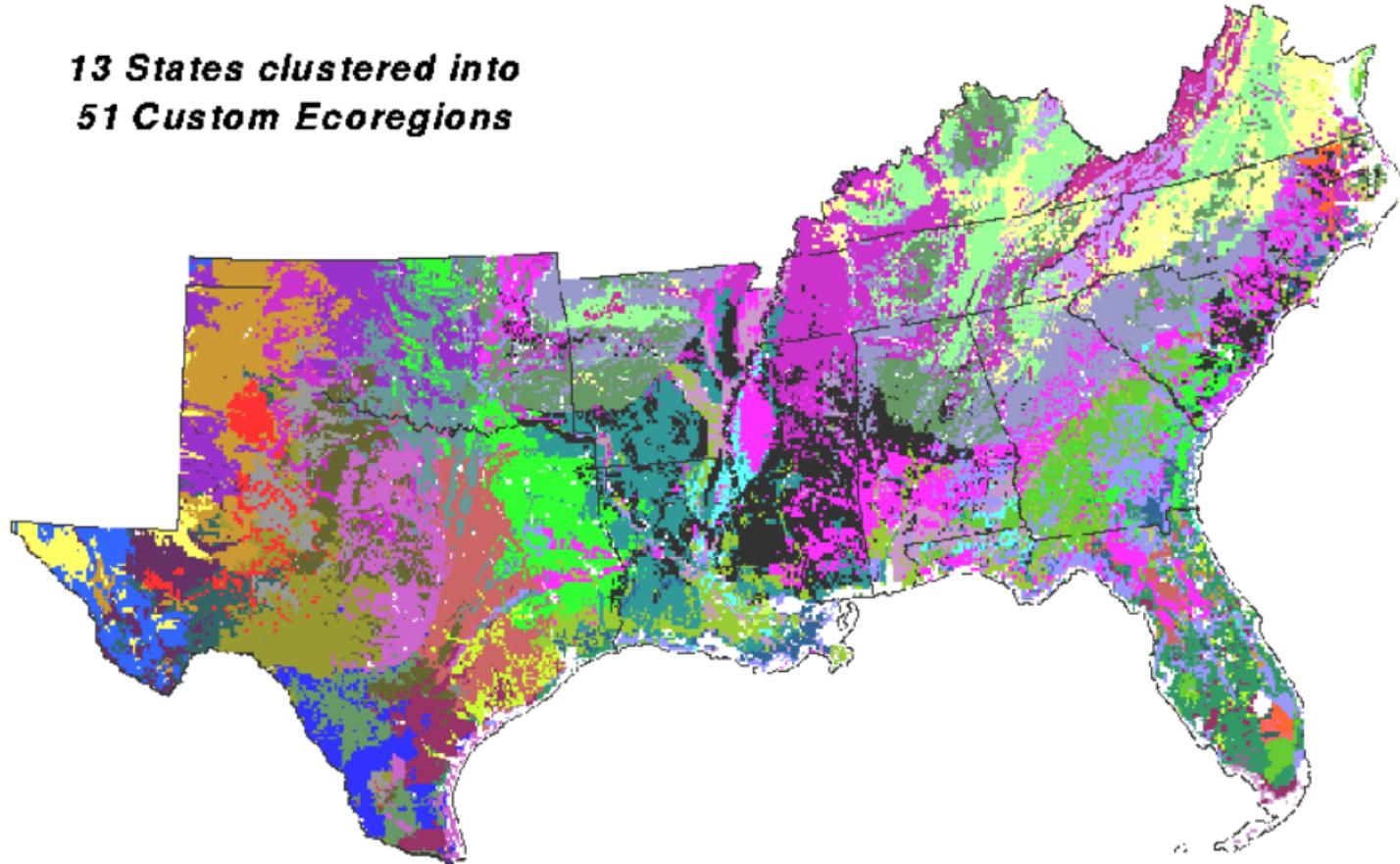


Estimated number of clusters: 3





*13 States clustered into
51 Custom Ecoregions*





Eddie Bell @ Lyst



Eddie Bell @ Lyst

Copyright 2016 Jeff Abrahamson, for private use by course students only

ML Week

23



Eddie Bell @ Lyst

Copyright 2016 Jeff Abrahamson, for private use by course students only

ML Week

24



Eddie Bell @ Lyst

Copyright 2016 Jeff Abrahamson, for private use by course students only

ML Week

25



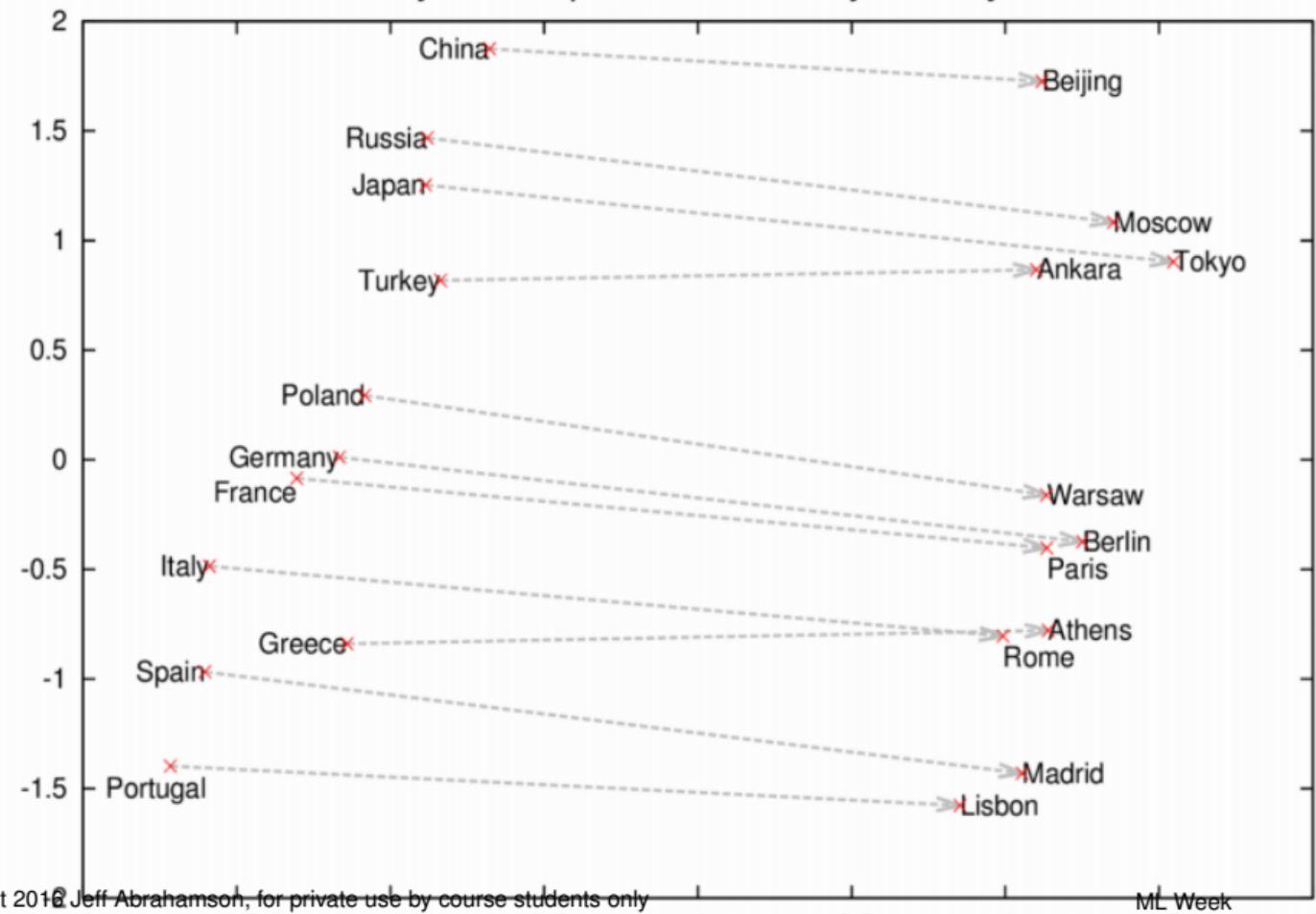
-1	0	+1
-2	0	+2
-1	0	+1



-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



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?

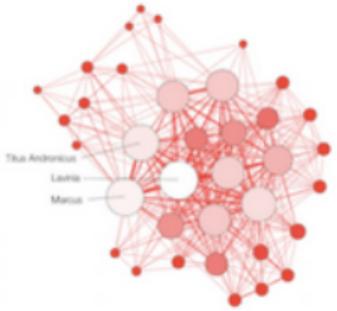


A bunch of bananas hanging from a tree

Google?

Copyright 2016 Jeff Abrahamson, for private use by course students only

ML Week



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



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



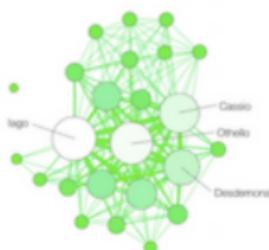
JULIUS CAESAR
Number of **46** | 34% Network density



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

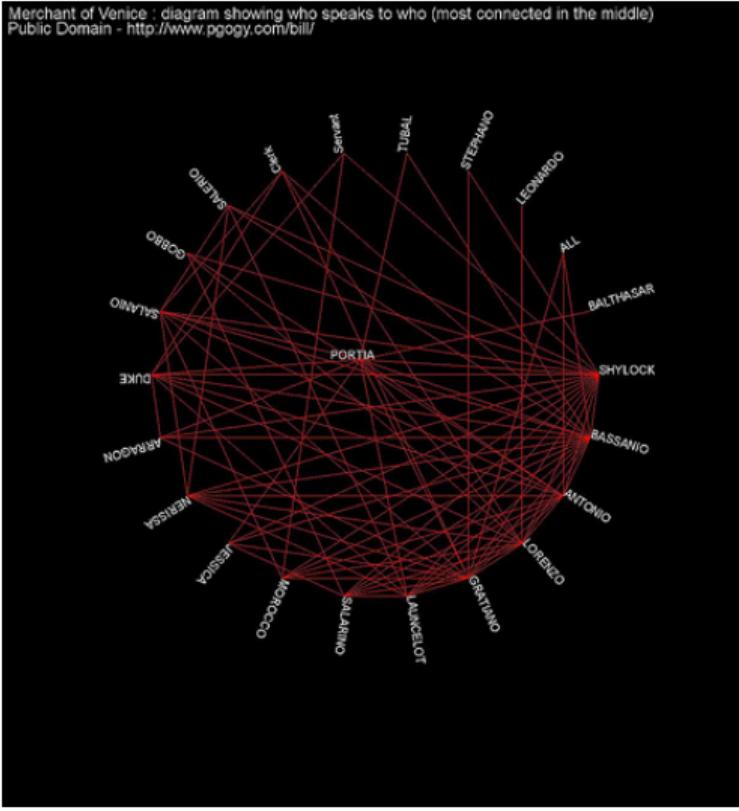


TROILUS AND CRESSIDA
Number of **35** | 40% Network density



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

Merchant of Venice : diagram showing who speaks to who (most connected in the middle)
Public Domain - <http://www.pgogy.com/bill/>



Friar Laurence: *No leisure required.*
 Friar: And I am nothing else to start his horse.
 Friar Laurence: *Unknown is the course, I like it not.*

And therefore have I little talked of love.

Friar Laurence: *Look, sir, here comes the lady towards my cell.*

Friar: *Pardon me, my lady and my wife.*
How now, my love, where comest thou?
 Friar: *From my round bower, where I did thee see.*
What courteous master!
 Friar Laurence: *That's a curious bairn.*

Friar Laurence: *Come to make confession to the father.*
To a priest that's wholly unblameable,
 Friar: *Be not sorry to hear that you have me.*
And confesse to you your thoughts,
 Friar: *For all I have seen, that you have hit.*
Being square behind your back, stand to your fate.
 Friar: *For such thy face is much about with me.*
The tears have got small victory by that,
 Friar: *How angry is man that wears all that green.*
And what I say, I say it to thy face,
 Friar: *Thou bairn art come, and thou hast deserved it.*
Or shall I come to you at evening mass.
 Friar Laurence: *My leisure serves me, gentle daughter, now.*
 Friar: *Juliet, on Thursday early will I rouse ye.*
Put hope, putchee, puthelp,
 Friar Laurence: *I hear thou must, and nothing may prorogue it.*

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

Friar Laurence: *Thou hast the strength of will to slay thyself.*

My dreams presage some joyful news at hand.

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

I will hence to-night.
I do beseech you, sir, have patience.
Tush, thou art deceived.

No import bid.

get thee gone.

Who calls me bid.

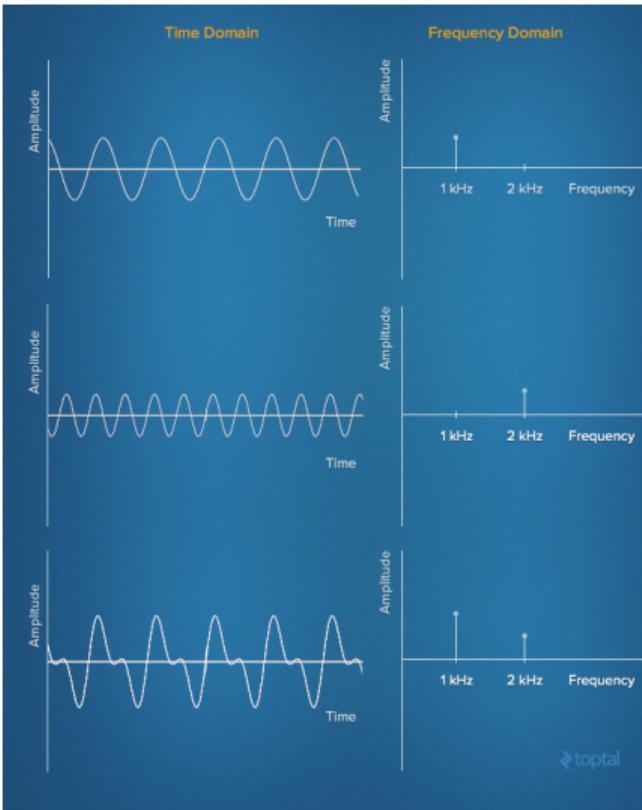
That the life-weary taker may fall dead.

But Maudlin's love.

famine is in thy cheeks.

No answer bid me bid no more.
Thy bosom is swelled by self.
Put this in any liquid thing you will.

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



<http://www.toptal.com/algorithms/>

shazam-it-music-processing-fingerprinting-and-recognition

Copyright 2016 Jeff Abrahamson, for private use by course students only

ML Week

34

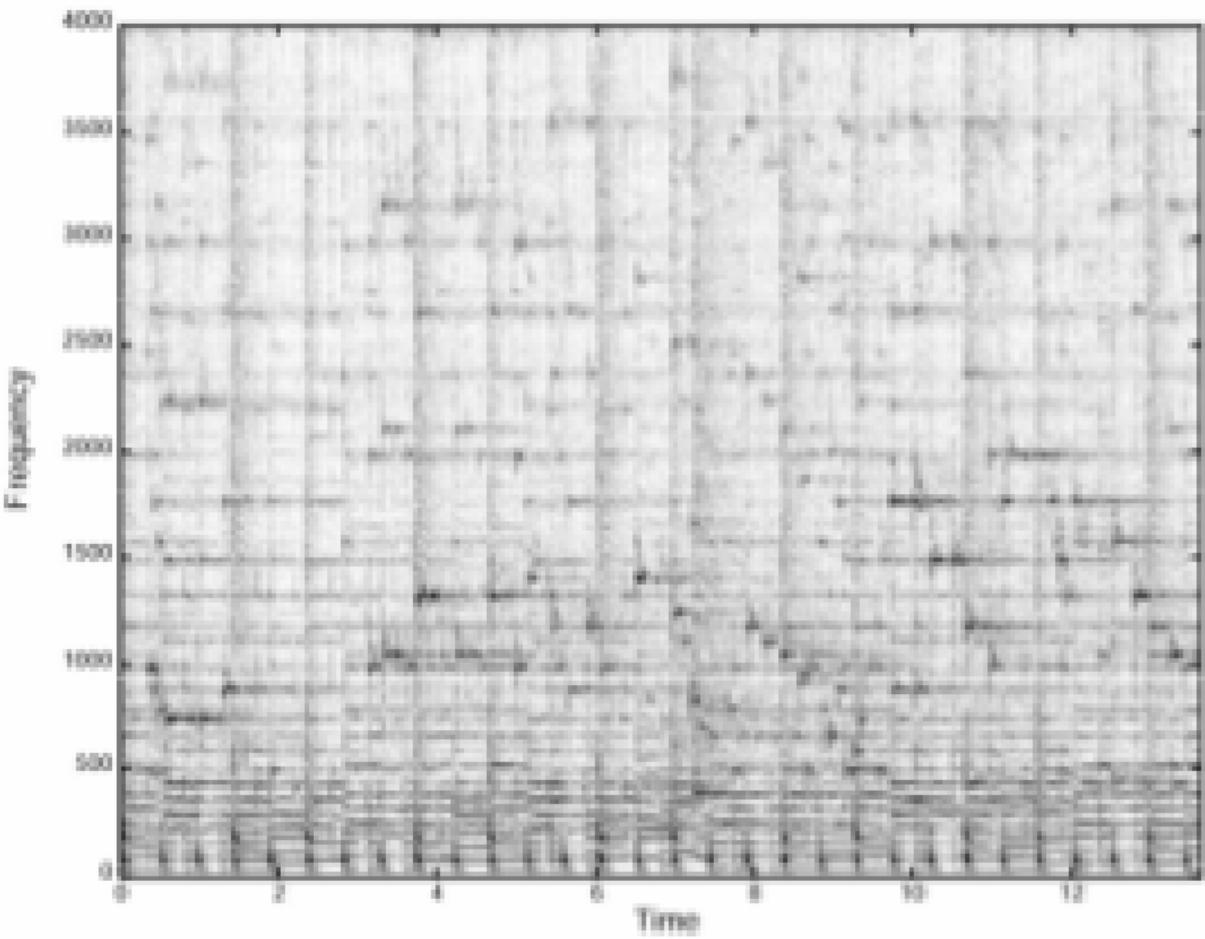


Fig 14 Spectrogram

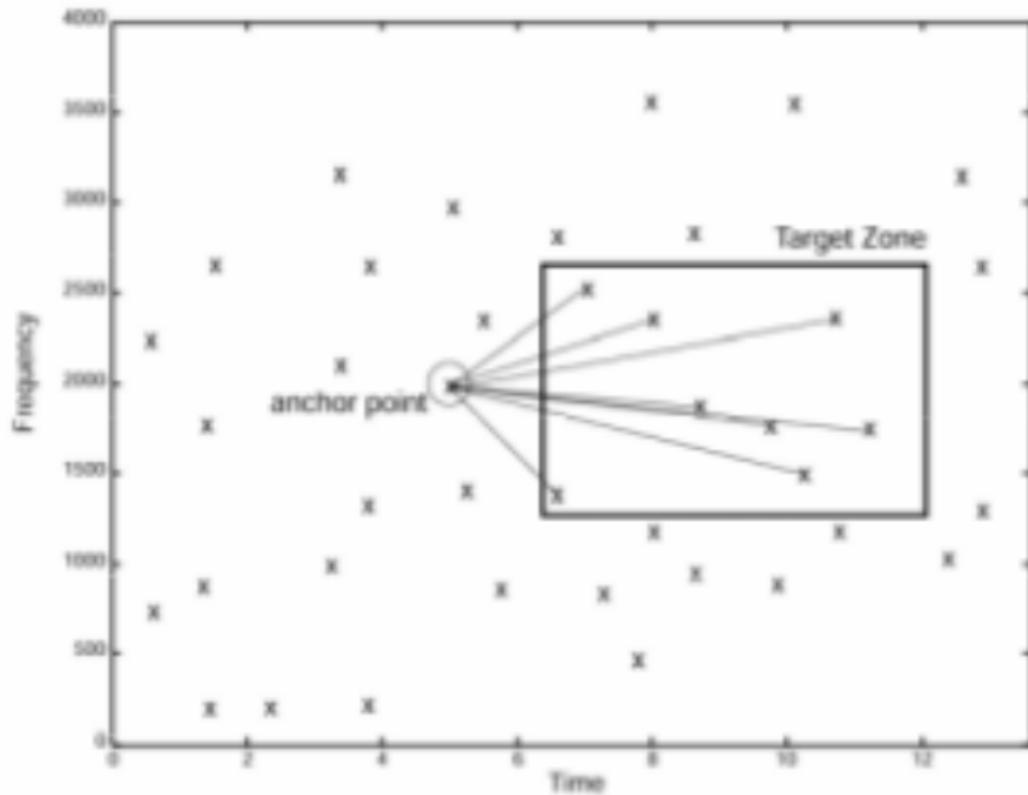


Fig. 1C - Combinatorial Hash Generation



Big data?

*[http://aadrake.com/
command-line-tools-can-be-235x-faster-than-your-hadoop-cluster.html](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?