

CORA JOHNSON-ROBERSON

BROWN UNIVERSITY

**COMPUTATIONAL ANALYSIS OF
MARGINALIZED CULTURAL PRODUCTION**

OVERVIEW

Why analyze?

Case Study: "Temporal and Regional Variation in Rap Lyrics" (2013)

Sampling of others' work

WHY COMPUTATIONAL ANALYSIS?

investigate concrete claims

counteract reductive narratives

reveal unnoticed relationships

DESIDERATA

sustainable

responsible

accessible to members of group

Let us build platforms that promote an understanding of the temporal vulnerability of the individual person or object; that more beautifully express the relationship of parts, one to another and to many a greater whole; and that instill, through depth of feeling in their users, an ethic of care—active, outward-facing, interdisciplinary, and expansive: sufficient to our daunting futures and broadened scope.

Bethany Nowviskie, “Capacity Through Care”

“TEMPORAL AND REGIONAL VARIATION IN RAP LYRICS”

Topic Modeling Workshop at Neural Information Processing Systems 2013

Impetus: interrogate claims in the literature on hip-hop, demonstrate the utility of techniques

DATA COLLECTION: OHHLA



~34,000 texts

user submitted (sometimes inconsistent)

DATA COLLECTION: ECHO NEST



Artist Locations

DATA COLLECTION: SPOTIFY



Album Release Dates

DATA PRE-PROCESSING: PYTHON

Python (Django database, ad-hoc scripts)

```
from django.db import models

class Song(models.Model):...
class Artist(models.Model):...
class Typist(models.Model):...
class Album(models.Model):...
class Place(models.Model):...
```

DATA PRE-PROCESSING: OPENREFINE



31722 records						
Show as: rows records			Show: 5 10 25 50 records			
☐ All	☐ _ - song	☐ _ - album	☐ _ - artist	☐ _ -	☐ _ -	☐ _ -
☐	☐	1.	Let's Get it On	Tension	10sion	Pittsbu
☐	☐	2.	Au Summum	Dans L'urgence	113	GhuldA
☐	☐	3.	Eye of the Storm	Consistency Theory	1200 Techniques	jgwasp
☐	☐	4.	Fork in the Road	Consistency Theory	1200 Techniques	Freedo
☐	☐	5.	Where Ur At?	Consistency Theory	1200 Techniques	newos_
☐	☐	6.	Karma	Infinite Styles EP	1200 Techniques	itzswift
☐	☐	7.	Handcuffs	Handcuffs (S)	12 O'Clock	pneum
☐	☐	8.	Nasty Immigrants	The Nutty Professor Soundtrack	12 O'Clock	OHHLA Flash
☐	☐	9.	Stay With Me	Ghost Dog soundtrack	Melodie	masta1
☐	☐	10.	Goin Back	1982	1982 (Statik Selektah & Termanology)	PDog1

METHOD: TOPIC MODELING

Topics

gene	0.04
dna	0.02
genetic	0.01
...	

life	0.02
evolve	0.01
organism	0.01
...	

brain	0.04
neuron	0.02
nerve	0.01
...	

data	0.02
number	0.02
computer	0.01
...	

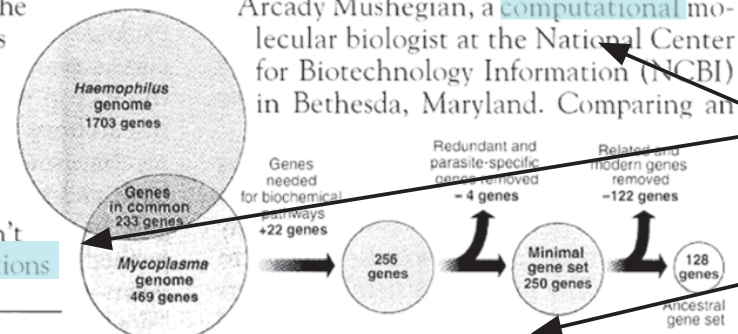
Documents

Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK—How many **genes** does an **organism** need to **survive**? Last week at the genome meeting here,* two genome researchers with radically different approaches presented complementary views of the basic genes needed for **life**. One research team, using **computer** analyses to compare known **genomes**, concluded that today's **organisms** can be sustained with just 250 genes, and that the earliest life forms required a mere 128 **genes**. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough.

Although the numbers don't match precisely, those **predictions**

"are not all that far apart," especially in comparison to the 75,000 **genes** in the human genome, notes Siv Andersson of Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a **genetic numbers** game, particularly as more and more **genomes** are completely mapped and sequenced. "It may be a way of organizing any newly **sequenced genome**," explains Arcady Mushegian, a **computational** molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an

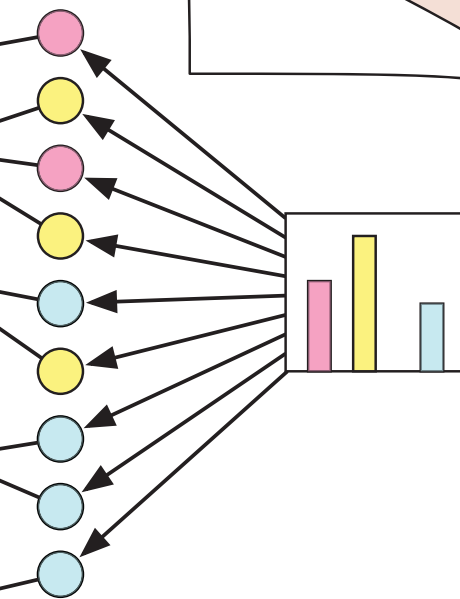


* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

Stripping down. **Computer analysis** yields an estimate of the minimum modern and ancient genomes.

SCIENCE • VOL. 272 • 24 MAY 1996

Topic proportions and assignments

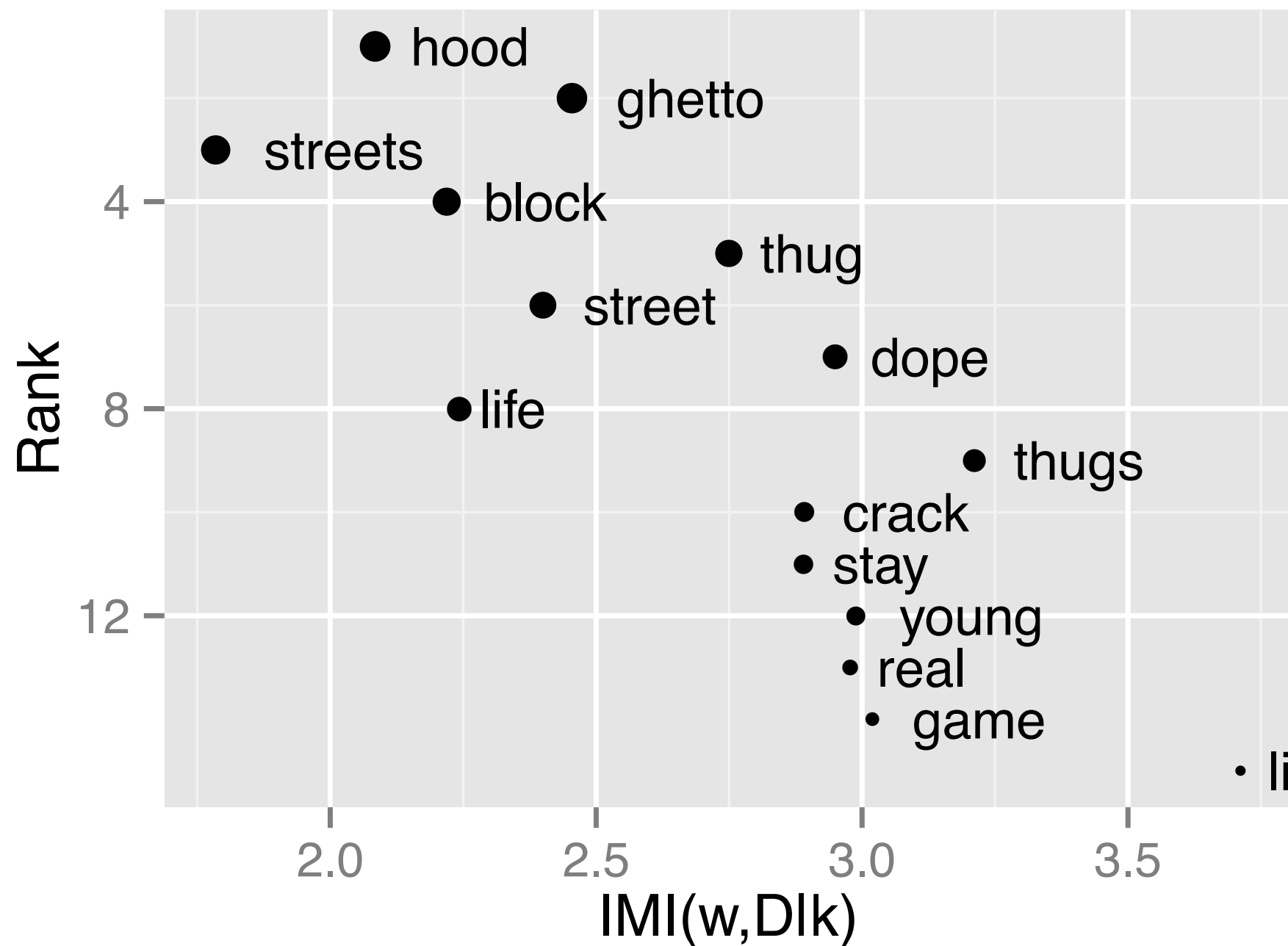


Blei, David M. "Probabilistic Topic Models."
Communications of the ACM 55, no. 4 (2012): 77-84.

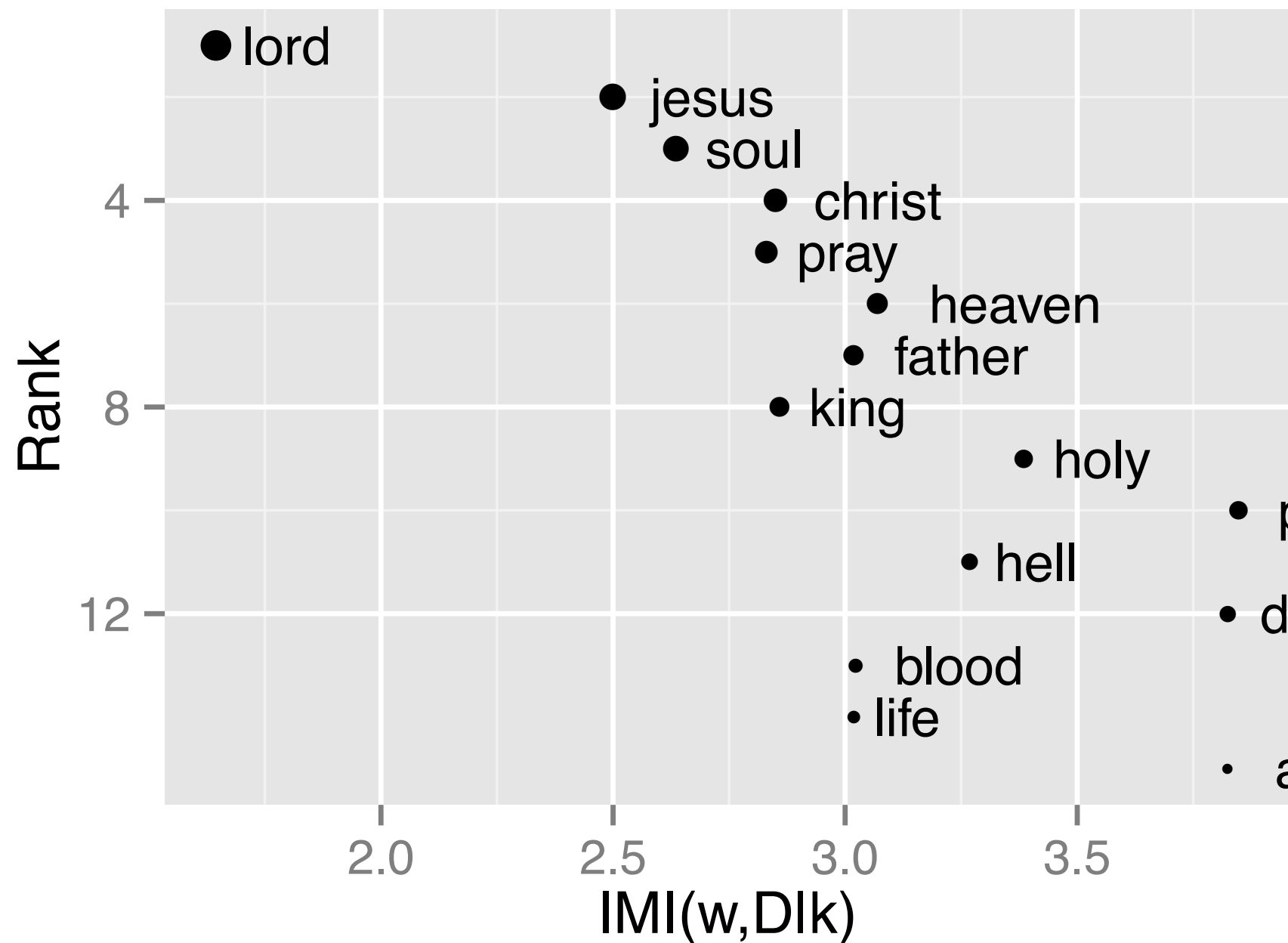
TOPIC MODELING

MALLET (via R)

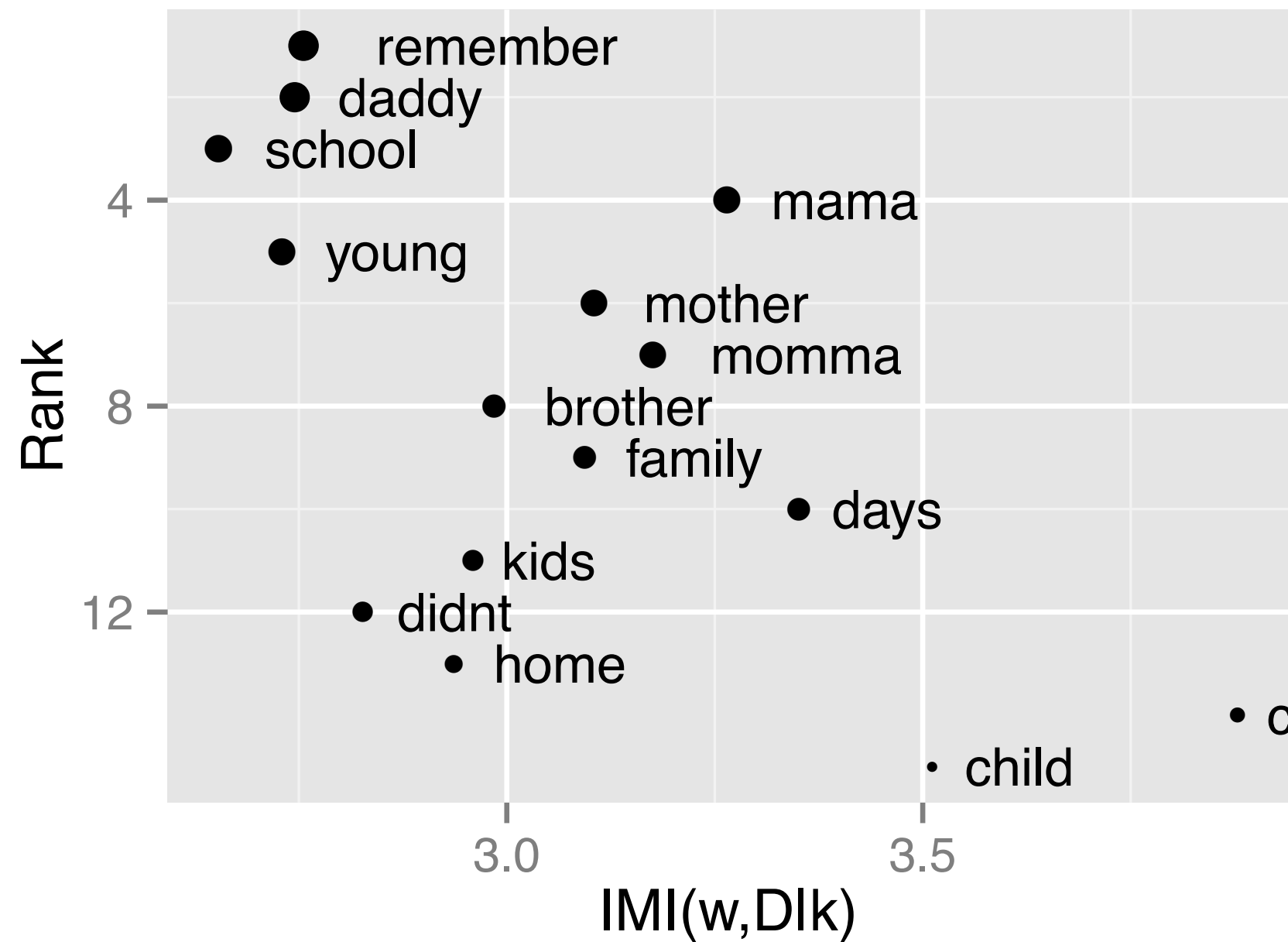
EXAMPLE TOPICS



EXAMPLE TOPICS



EXAMPLE TOPICS



EXAMPLE TEXT: 2PAC'S "DEAR MAMA"

When I was young me and my mama had beef

Seventeen years old kicked out on the streets

Though back at the time, I never thought I'd see her face

Ain't a woman alive that could take my mama's place

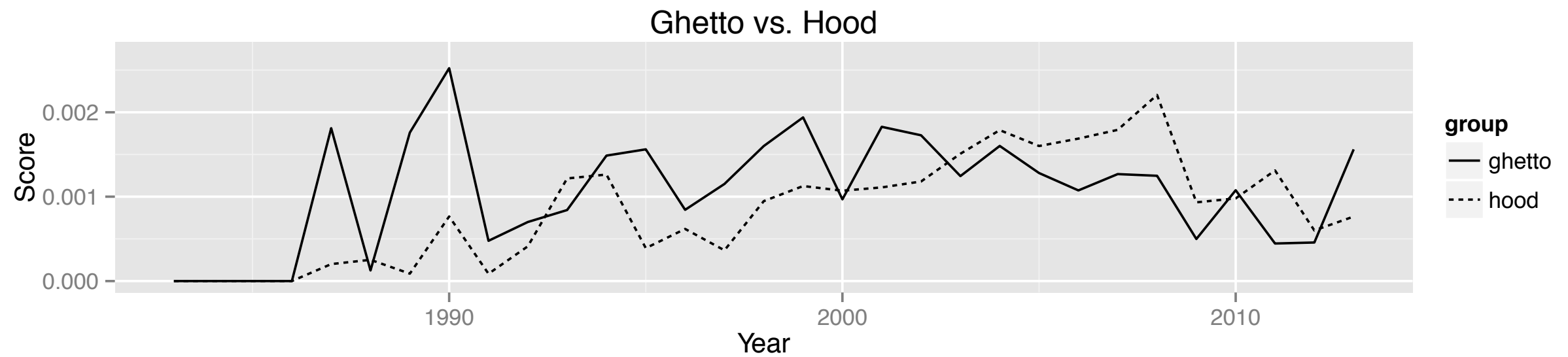
Suspended from school; and scared to go home, I was a fool

with the big boys, breakin all the rules

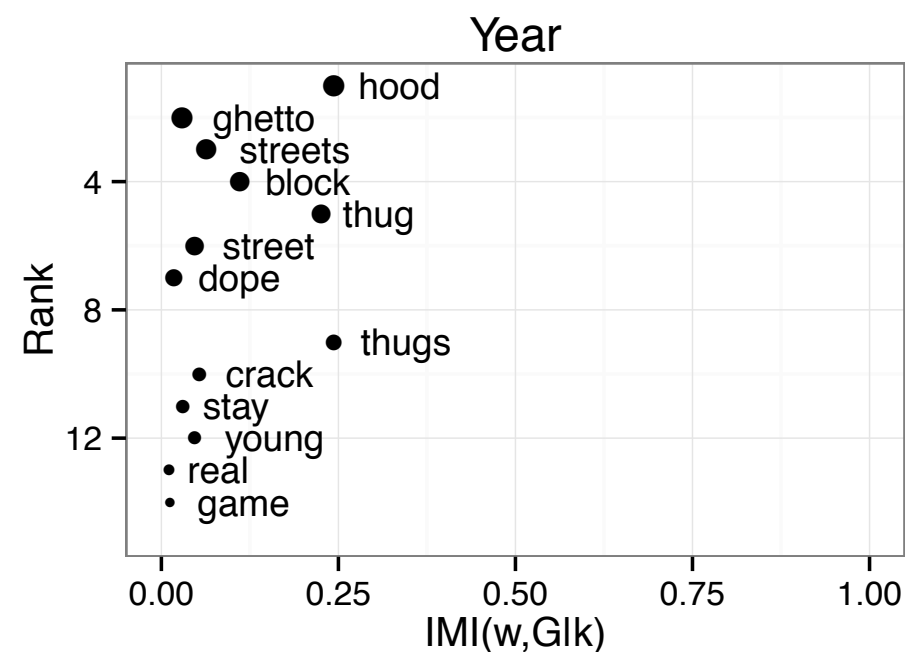
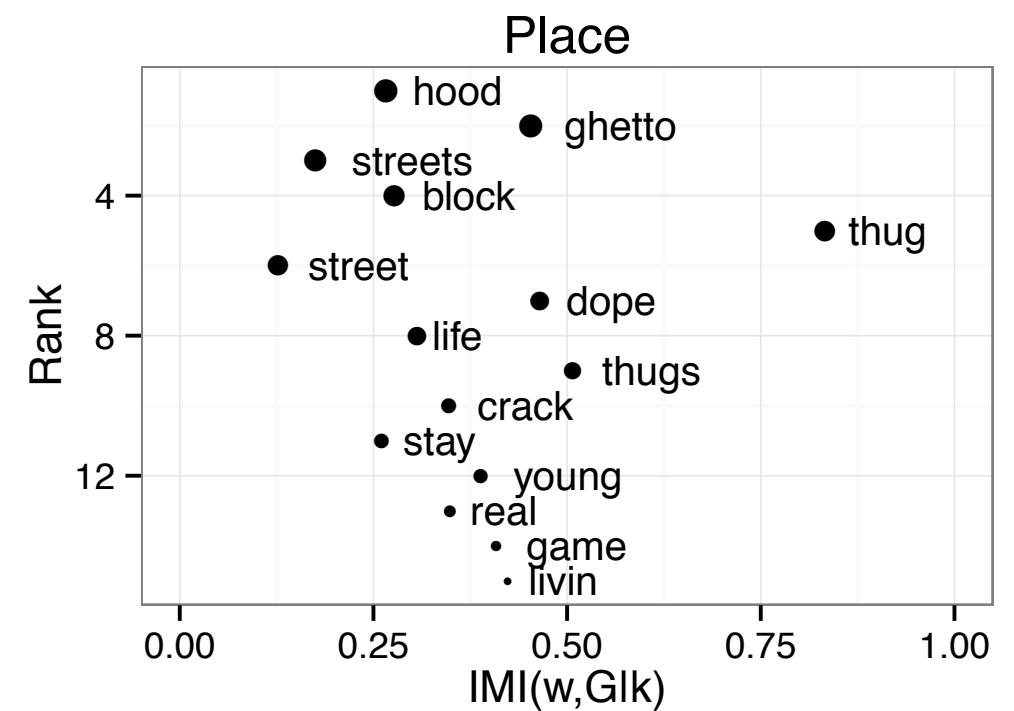
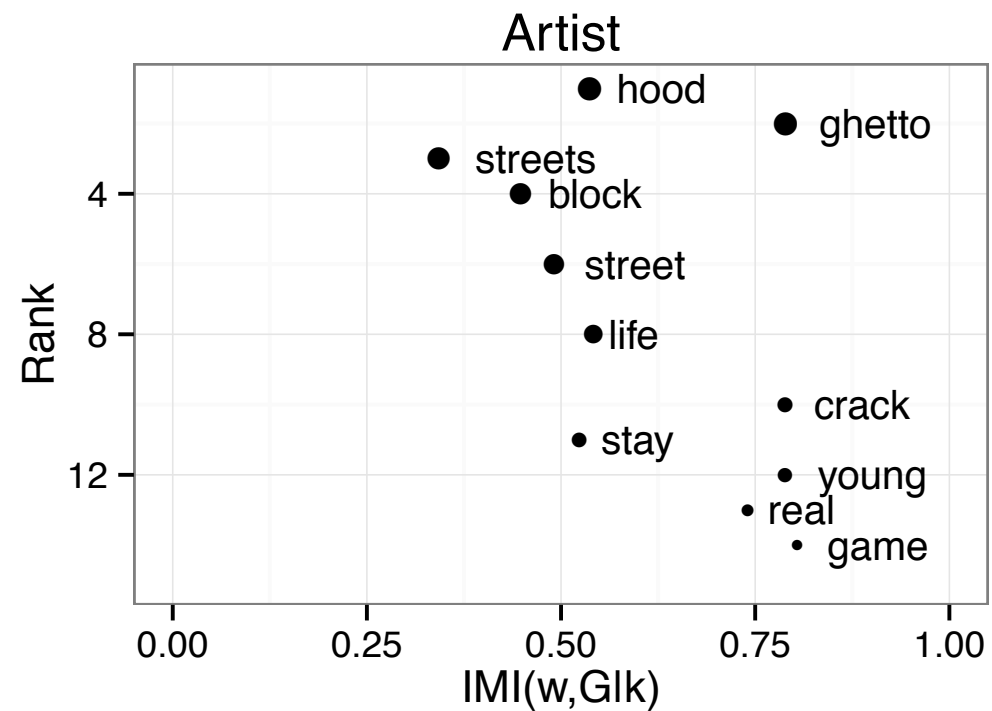
I shed tears with my baby sister

Over the years we was poorer than the other little kids

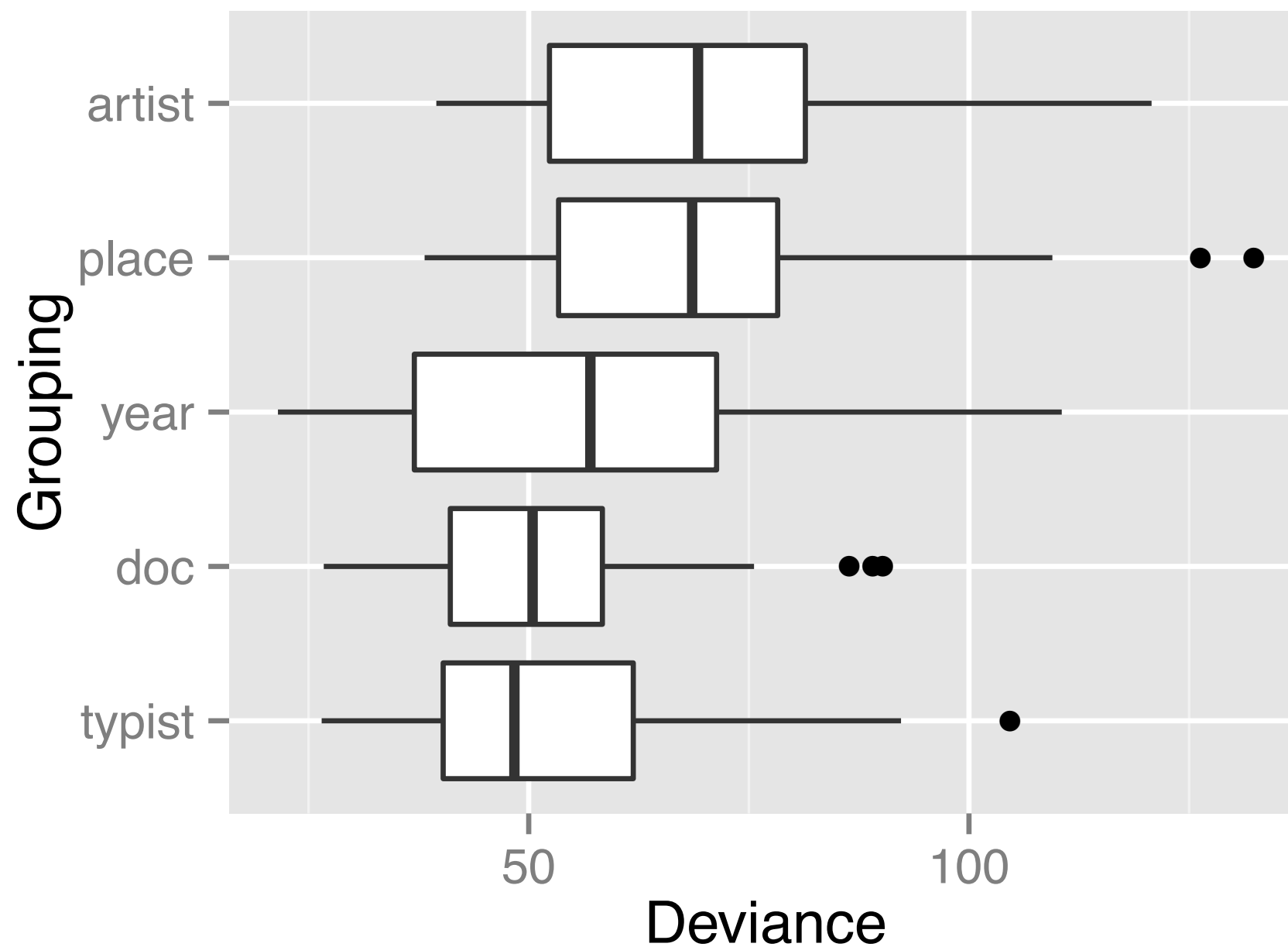
"GHETTO" VS. "HOOD"



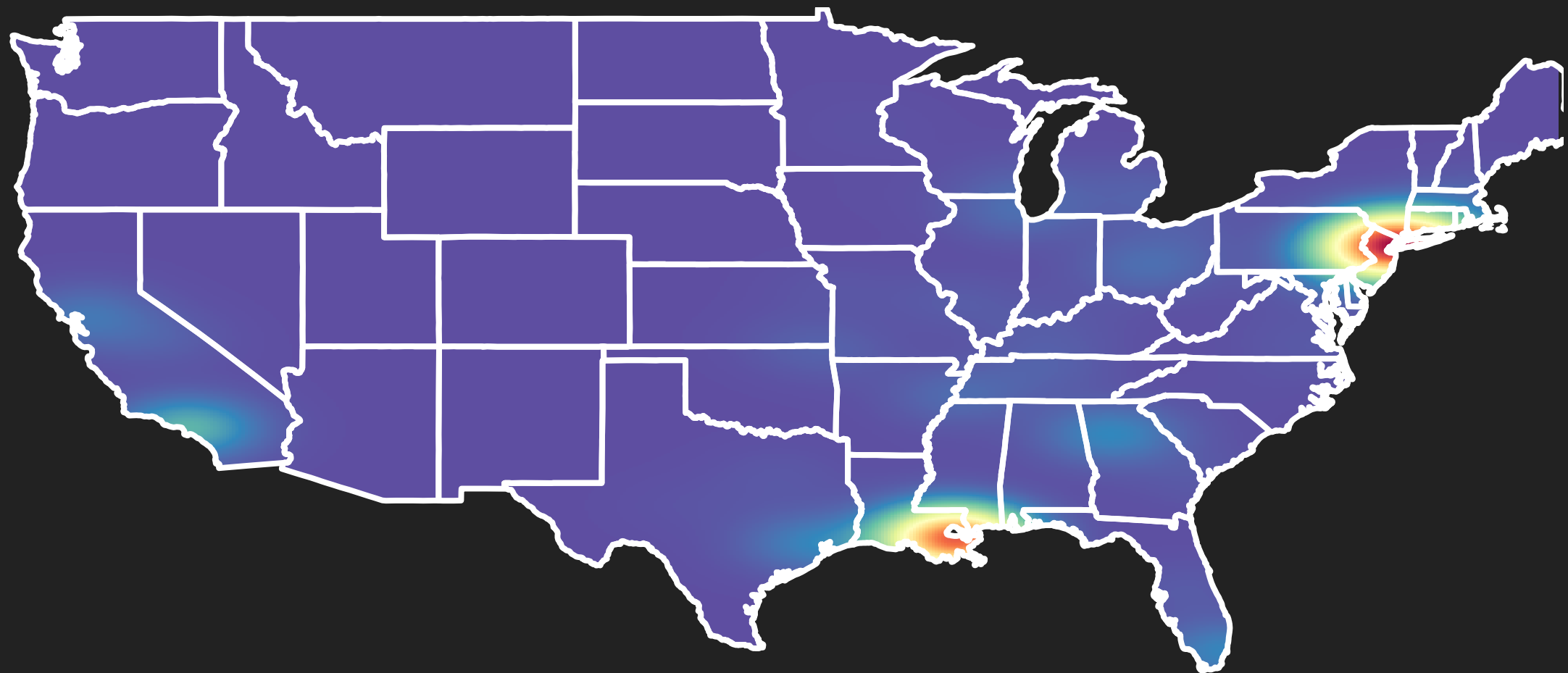
"GHETTO" VS. "HOOD"



SYSTEMATIC DEVIATIONS DUE TO METADATA



REGIONAL DISTRIBUTION: {GHETTO, THUG, HOOD}

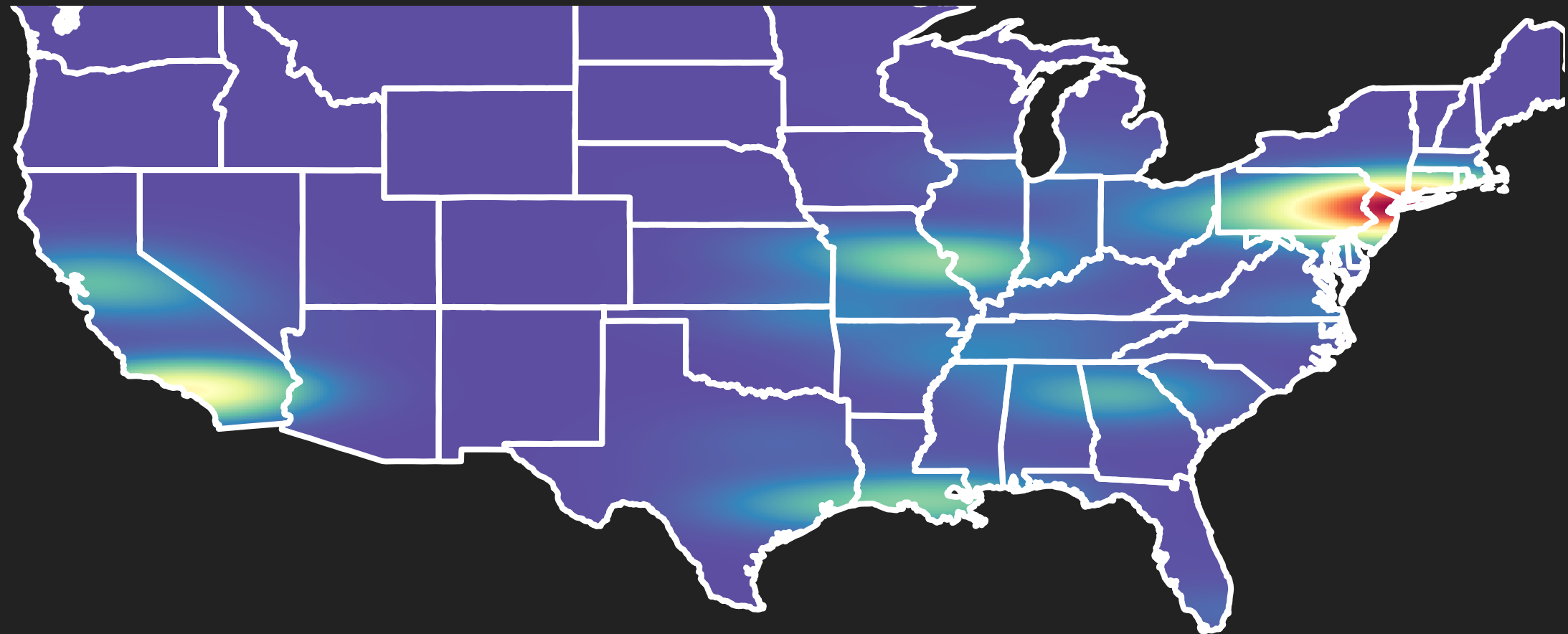


Fewer Occurrences



More Occurrences

REGIONAL DISTRIBUTION: {GANGSTA, GANG, GANGSTAS}



Fewer Occurrences



More Occurrences

AFTERLIFE

this talk!

the odd e-mail (mostly from students)

much more to be done

OTHERS HAVE WORKED ON...

Hip-hop

Slave Narratives

Black Women's Suffrage

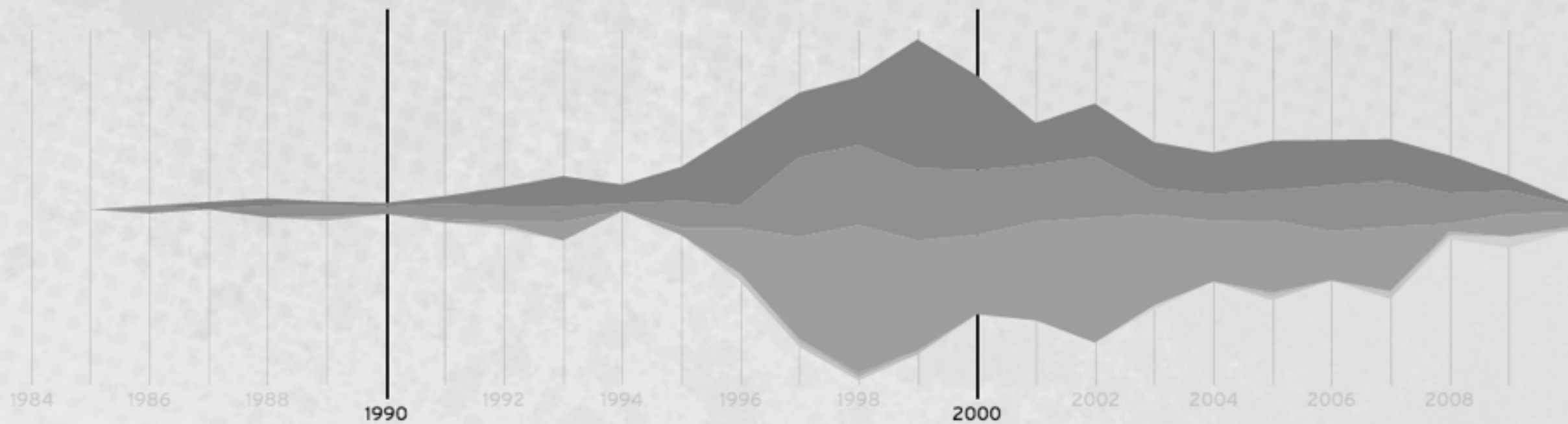
Black Press

TAHIR HEMPHILL

HIP HOP WORD COUNT

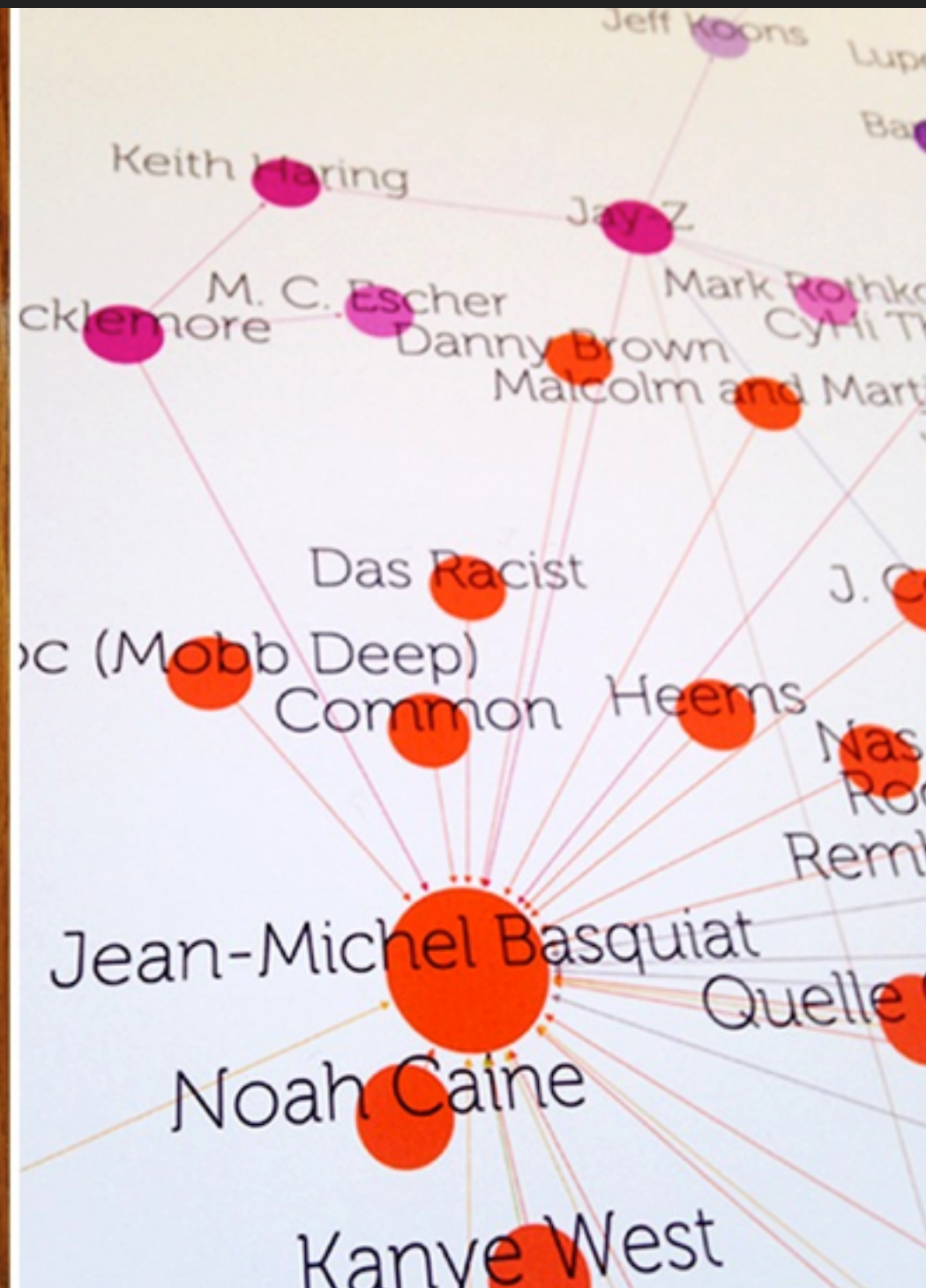
Brand Mentions 1980 - 2010

CHAMPAGNE

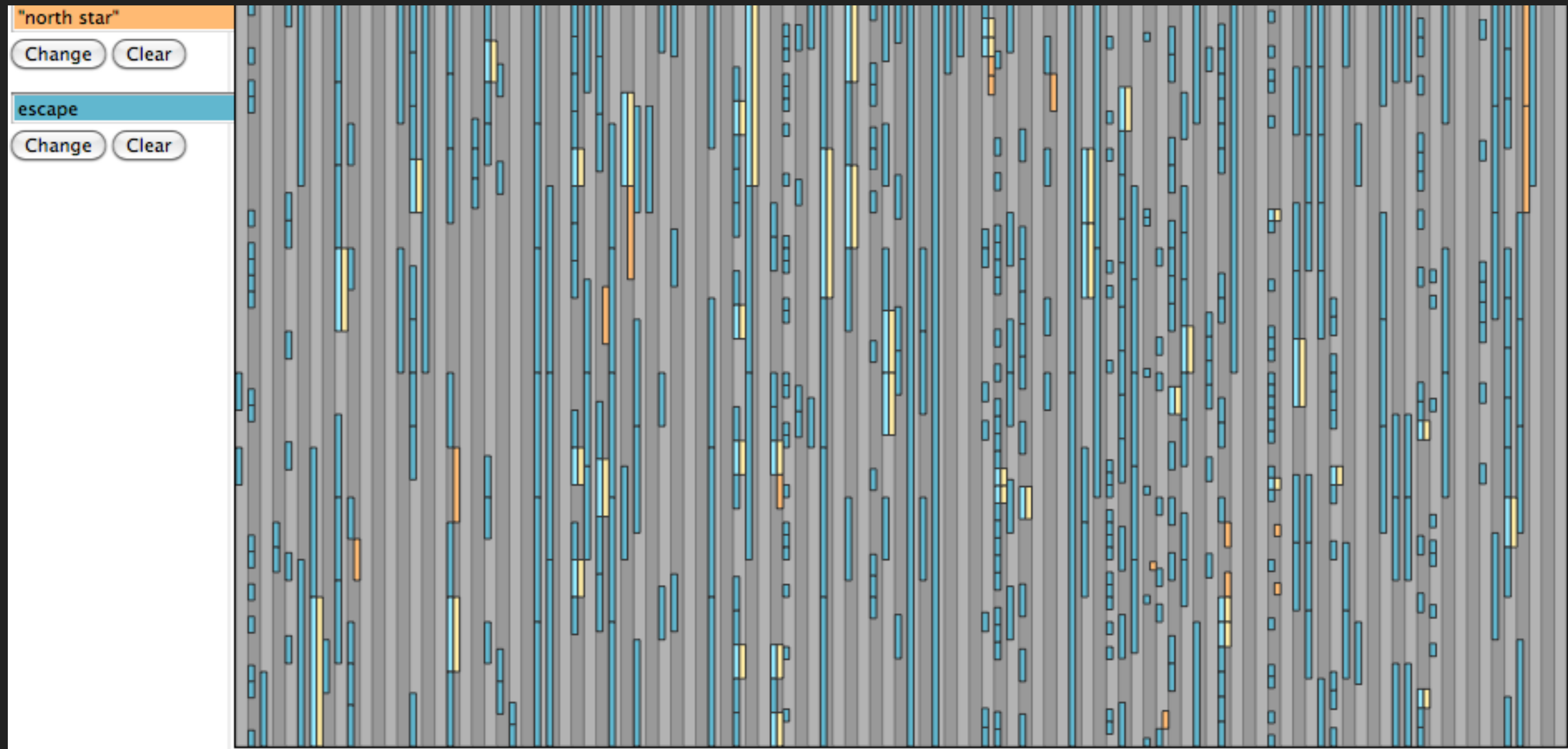


1980 - 2010

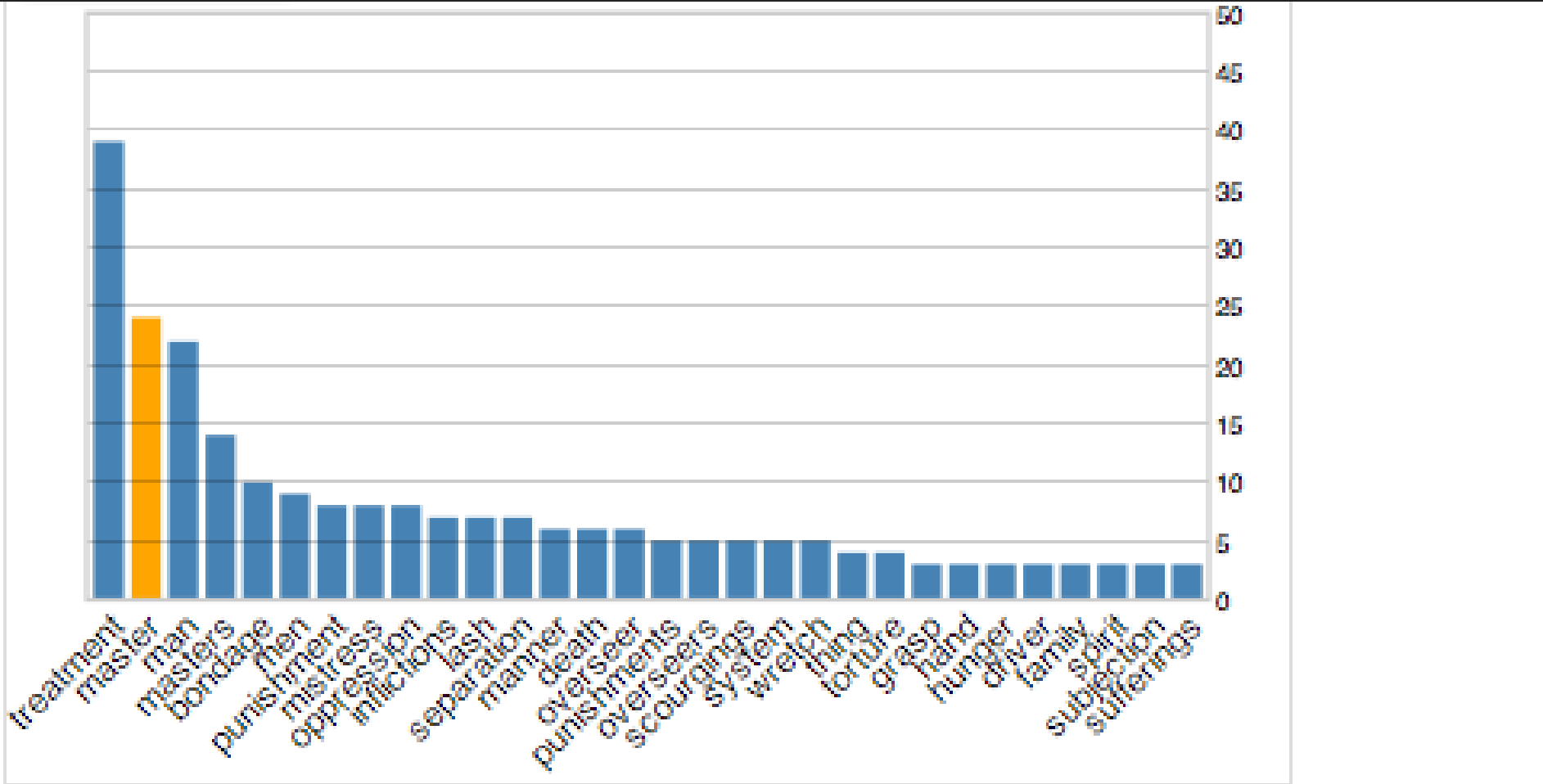
TAHIR HEMPHILL



ADITI MURALIDHARAN: WORDSEER (2010–)



ADITI MURALIDHARAN: WORDSEER (2010-)



24 results

View	Sentence	Type	Title	Author	Date	Publisher
	On my knees , with tears in my eyes , with terror in my countenance , and fervency in all my features , I implored Mr. Ballard to buy me , but he again refused , and I was taken back to my dreaded cruel master .	section	Uncle Tom's Companions: Or, Facts Stranger Than Fiction. A Supplement to Uncle Tom's Cabin: Being Startling Incidents in the Lives of Celebrated Fugitive Slaves.	Edwards, John Passmore, 1823-1911	1852	EDWARDS AND CO., 2, HORSESHOE COURT LUDGATE HILL.

MICHELLE MORAVEC: BLACK WOMAN SUFFRAGE DATABASE

her _____ in subcorpora Black Woman Suffragists Database and History of Woman Suffrage

BWSD

#Total No. of Cluster Types: 280			
#Total No. of Cluster Tokens: 440			
1	14	13	her hands ←
2	13	6	her own
3	8	4	her child
4	8	5	her hand ←
5	7	6	her to
6	6	5	her a
7	6	5	her children
8	6	5	her heart ←
9	6	4	her husband
10	6	5	her with
11	5	4	her race
12	4	3	her back ←
13	4	4	her from
14	4	4	her in
15	4	4	her life
16	4	4	her work
17	4	4	her, and
18	3	2	her address
19	3	3	her and
20	3	2	her arms ←
21	3	3	her bosom ←
22	3	3	her feet ←
23	3	3	her friends
24	3	2	her head ←
25	3	2	her letter

#Total No. of Cluster Types: 538			
#Total No. of Cluster Tokens: 1069			
1	77	32	her own
2	36	15	her husband
3	35	22	her to
4	19	14	her children
5	17	15	her rights
6	15	13	her right
7	14	11	her as
8	11	8	her in
9	11	9	her the
10	8	7	her child
11	8	8	her for
12	8	5	her home
13	8	8	her life
14	7	4	her mind
15	7	7	her property
16	7	3	her sphere
17	6	4	her by
18	6	5	her highest
19	6	4	her on
20	6	4	her sister
21	6	4	her true
22	5	4	her a
23	5	4	her duties
24	5	5	her existence
25	5	4	her head

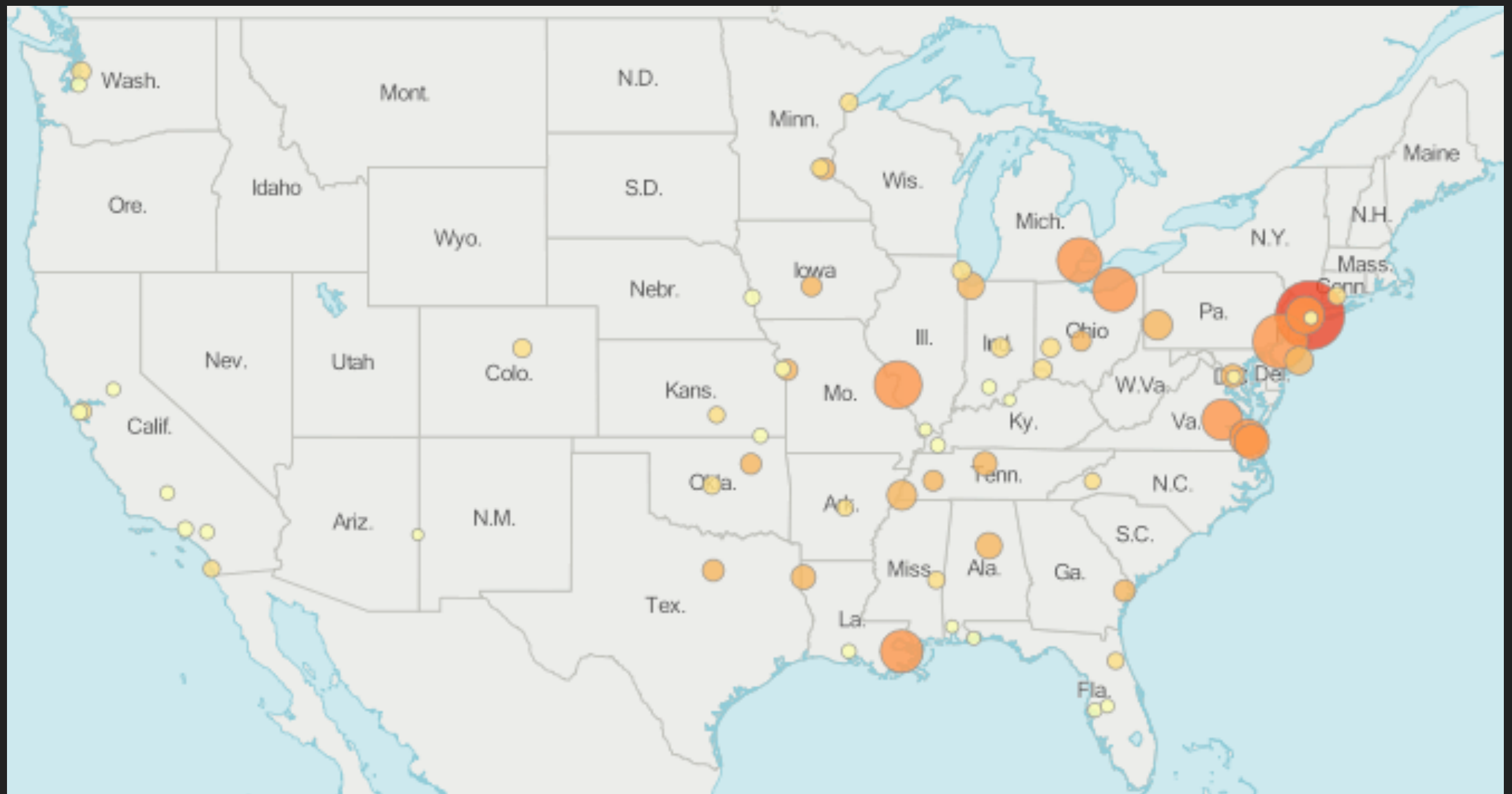
HWS

As Hazel Carby notes, [Frances] Harper crafted a discourse for black women that “address[ed] exclusion from the ideology of true womanhood” and “rescue[d] their bodies from a persistent association with illicit sexuality.” That she did this by relying in part on embodied references drawn from sentimental conventions makes perfect sense.

Michelle Moravec

KIM GALLON: BLACK PRESS RESEARCH COLLECTIVE

Chicago Defender's Standing Dealers, 1919



WHY AREN'T THERE MORE?

archiving & digitization challenges

skepticism?

lack of resources?

CALLS TO ACTION

lift up work by/for marginalized groups

open up data and process when possible

Thanks! Q/A

@cora_jr

me@corajr.com