Lexicocalorimeter: Real-time health measurement

Last updated: 2023/08/22, 11:48:21 EDT

Principles of Complex Systems, Vols. 1, 2, & 3D
CSYS/MATH 6701, 6713, & a pretend number,
2023–2024 | @pocsvox

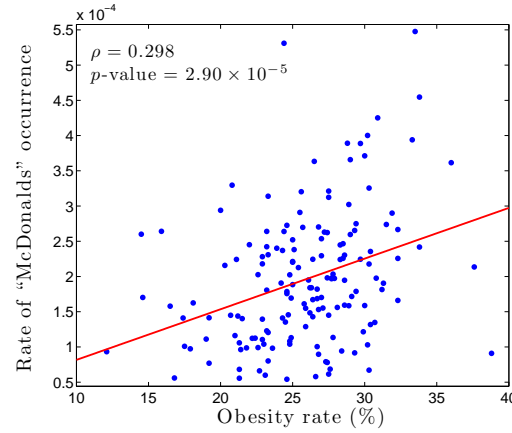### Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Santa Fe Institute | University of Vermont

The PoCSverse
Lexicocalorimeter
1 of 27
The LCM
Other lexical meters
References

Outline

The LCM

Other lexical meters

References

The PoCSverse
Lexicocalorimeter
2 of 27
The LCM
Other lexical meters
References

Panometer.org
Measuring everything.
Coming soon

Obesity and tweets—"McDonalds":

$\rho = 0.298$
$p\text{-value} = 2.90 \times 10^{-5}$

Obesity and tweets—"Brunch":

$\rho = -0.372$
$p\text{-value} = 1.24 \times 10^{-7}$

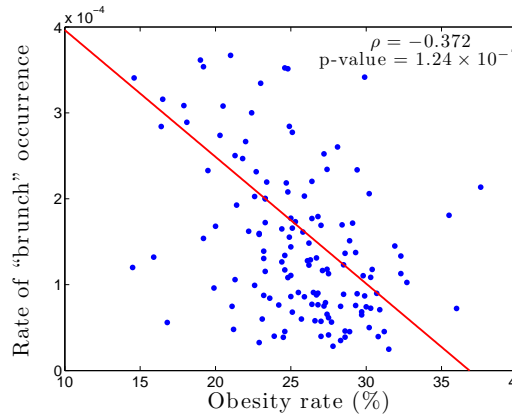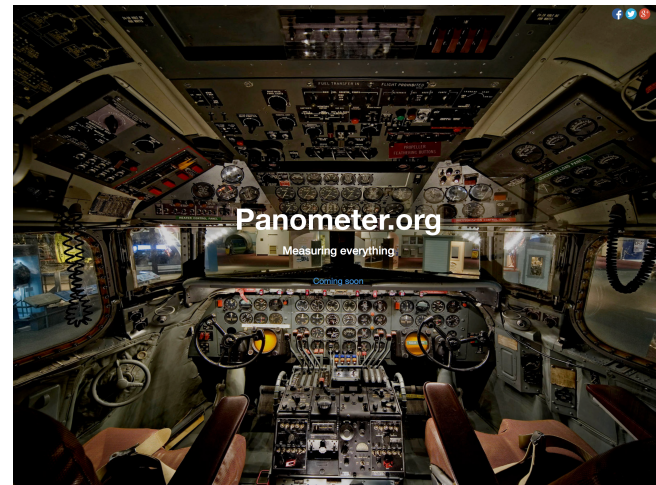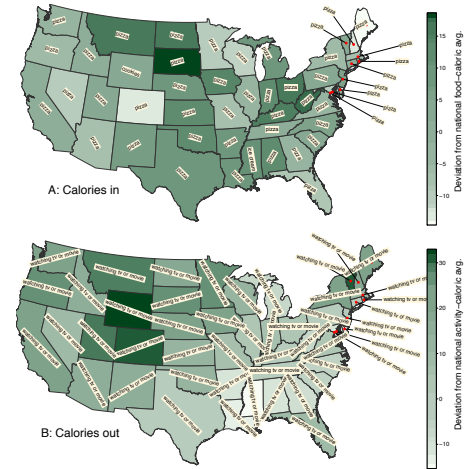The PoCSverse
Lexicocalorimeter
4 of 27
The LCM
Other lexical meters
References

The PoCSverse
Lexicocalorimeter
5 of 27
The LCM
Other lexical meters
References

Word	ρ	p-value
cafe	-0.509	6.07×10^{-14}
sushi	-0.487	9.93×10^{-13}
brewery	-0.469	8.67×10^{-12}
restaurant	-0.448	8.93×10^{-11}
bar	-0.435	3.59×10^{-10}
banana	-0.434	3.77×10^{-10}
apple	-0.408	5.22×10^{-9}
fondue	-0.403	8.34×10^{-9}
wine	-0.400	1.08×10^{-8}
delicious	-0.392	2.17×10^{-8}
dinner	-0.386	3.85×10^{-8}
coffee	-0.384	4.51×10^{-8}
bakery	-0.383	5.12×10^{-8}
bean	-0.378	7.88×10^{-8}
espresso	-0.377	8.47×10^{-8}
cuisine	-0.376	8.82×10^{-8}
foods	-0.374	1.07×10^{-7}
tofu	-0.372	1.27×10^{-7}
brunch	-0.368	1.79×10^{-7}
veggie	-0.364	2.46×10^{-7}
organic	-0.361	3.13×10^{-7}
booze	-0.360	3.34×10^{-7}
grill	-0.354	5.4×10^{-7}
chocolate	-0.351	6.77×10^{-7}
#vegan	-0.350	7.47×10^{-7}

mcdonalds	0.246	6.18×10^{-4}
eat	0.241	8.22×10^{-4}
wings	0.222	2.13×10^{-3}
hungry	0.210	3.65×10^{-3}
heartburn	0.194	7.37×10^{-3}
ham	0.177	1.45×10^{-2}

The PoCSverse
Lexicocalorimeter
6 of 27
The LCM
Other lexical meters
References

---

http://panometer.org/instruments/lexicocalorimeter/

The PoCSverse
Lexicocalorimeter
7 of 27
The LCM
Other lexical meters
References

The Lexicocalorimeter:

A: Calories in

B: Calories out

The Lexicocalorimeter:

A: Calories in

B: Calories out

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
10 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
13 of 27

The LCM
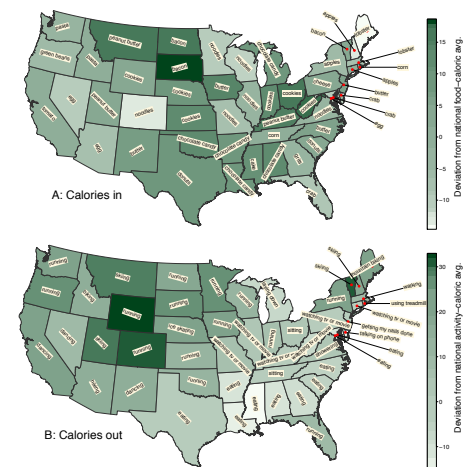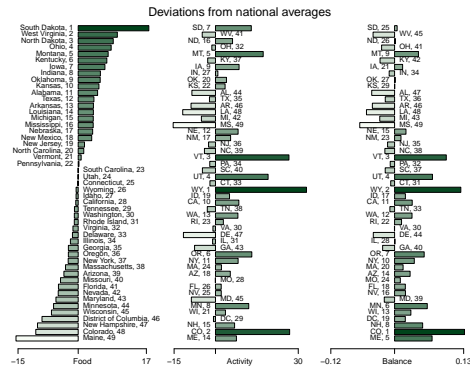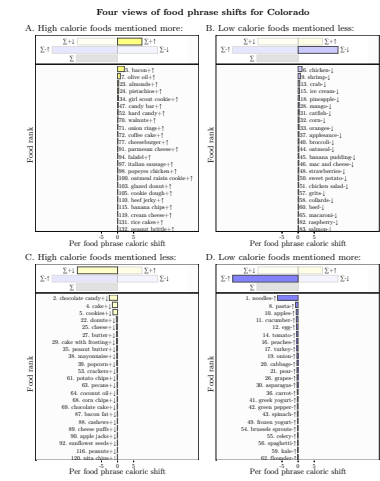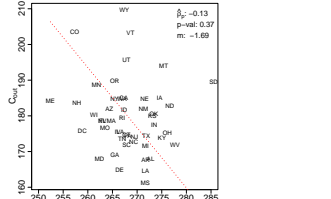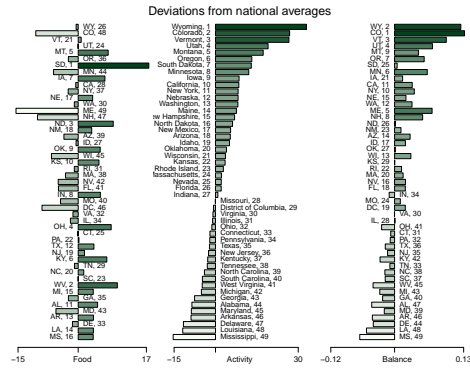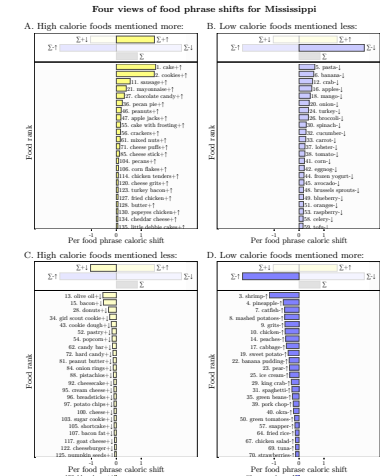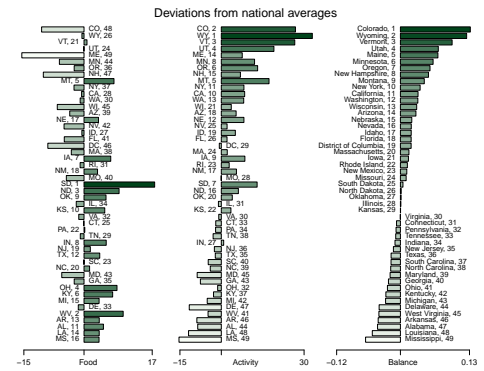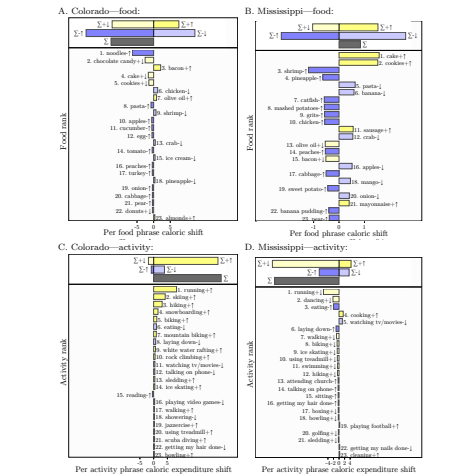
Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
16 of 27

The LCM

Other lexical
meters

References

Deviations from national averages

Four views of food phrase shifts for Colorado

A. High calorie foods mentioned more:
B. Low calorie foods mentioned less:
C. High calorie foods mentioned less:
D. Low calorie foods mentioned more:

Per food phrase caloric shift

The Lexicocalorimeter:
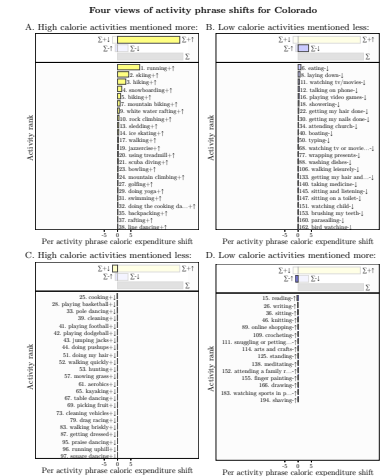
The PoCSverse
Lexicocalorimeter
11 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
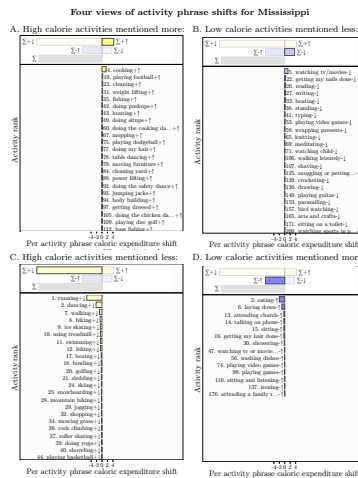Lexicocalorimeter
14 of 27
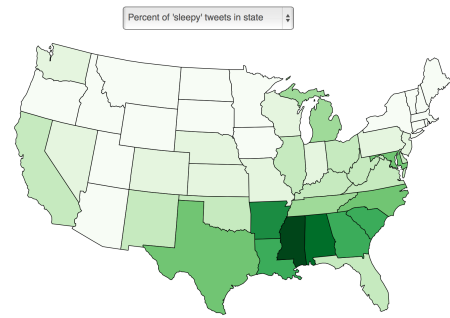
The LCM

Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
17 of 27

The LCM

Other lexical
meters

References

Deviations from national averages

Four views of food phrase shifts for Mississippi

A. High calorie foods mentioned more:
B. Low calorie foods mentioned less:
C. High calorie foods mentioned less:
D. Low calorie foods mentioned more:

Per food phrase caloric shift

The Lexicocalorimeter:

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
12 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
15 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

The PoCSverse
Lexicocalorimeter
18 of 27

The LCM

Other lexical
meters

References

Deviations from national averages

A. Colorado—food:
B. Mississippi—food:
C. Colorado—activity:
D. Mississippi—activity:

Per food phrase caloric shift
Per activity phrase caloric expenditure shift

Four views of activity phrase shifts for Colorado

A. High calorie activities mentioned more:
B. Low calorie activities mentioned less:
C. High calorie activities mentioned less:
D. Low calorie activities mentioned more:

Per activity phrase caloric expenditure shift

The Lexicocalorimeter:

Four views of activity phrase shifts for Mississippi

A. High calorie activities mentioned more:

B. Low calorie activities mentioned less:

C. High calorie activities mentioned less:

D. Low calorie activities mentioned more:

Per activity phrase caloric expenditure shift

The PoCSverse
Lexicocalorimeter
19 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

	ρ̂s for C_{out}	q-val	ρ̂s for C_{in}	q-val	ρ̂s for C_{out}
Physical Inactivity	−0.78	q: 2.5 × 10⁻⁹			
High Blood Pressure	−0.77	q: 3.5 × 10⁻⁹			
Diabetes	−0.74	q: 1.4 × 10⁻⁸			
Overweight	−0.74	q: 1.6 × 10⁻⁸			
Heart Disease Death	−0.73	q: 2.8 × 10⁻⁸			
Life Expectancy	0.68	q: 5.7 × 10⁻⁷			

Scatter plots versus Caloric Balance:
- Physical Inactivity: $\hat{\rho}_s: -0.78$, $q: 2.5 \times 10^{-9}$
- High Blood Pressure: $\hat{\rho}_s: -0.77$, $q: 3.5 \times 10^{-9}$
- Diabetes: $\hat{\rho}_s: -0.74$, $q: 1.4 \times 10^{-8}$
- Overweight: $\hat{\rho}_s: -0.74$, $q: 1.6 \times 10^{-8}$
- Heart Disease Death: $\hat{\rho}_s: -0.73$, $q: 2.8 \times 10^{-8}$
- Life Expectancy: $\hat{\rho}_s: 0.68$, $q: 5.7 \times 10^{-7}$

x-axis: Caloric Balance

The Lexicocalorimeter:

TABLE I. Spearman correlation coefficients, $\hat{\rho}_s$, and Benjamini-Hochberg q-values for caloric input C_{in}, caloric output C_{out}, and caloric ratio $C_{rat} = C_{out}/C_{in}$, and demographic, data related to food and physical activity, Big Five personality traits [31], health and well-being rankings by state, and socioeconomic status, correlated, ordered from strongest to weakest Spearman correlations with caloric ratio. Two breaks in the table indicate significance levels of 0.01 and 0.05 for the Benjamini-Hochberg q of C_{rat}; corresponding to the first 24 health and/or well-being quantities and then the next four, numbers 25 to 28. The bottom 9 quantities were not significantly correlated with C_{rat} according to our tests. Tables S1, S2, and S3 present the same analysis for caloric measures including phrases representing liquids, and for the difference $C_{diff}(\alpha) = \alpha C_{out} - (1 - \alpha)C_{in}$, both without and with liquids included.

The PoCSverse
Lexicocalorimeter
21 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

TABLE S1. Identical to Tab. 1 but with liquids included. Spearman correlation coefficients, $\hat{\rho}_s$, and Benjamini-Hochberg q-values for caloric input C_{in}, caloric output C_{out}, and caloric ratio $C_{rat} = C_{out}/C_{in}$ and demographic data related to food and physical activity, Big Five personality traits [31], health and well-being rankings by state, and socioeconomic status, correlated, ordered from strongest to weakest Spearman correlations with caloric ratio.

The PoCSverse
Lexicocalorimeter
22 of 27

The LCM

Other lexical
meters

References

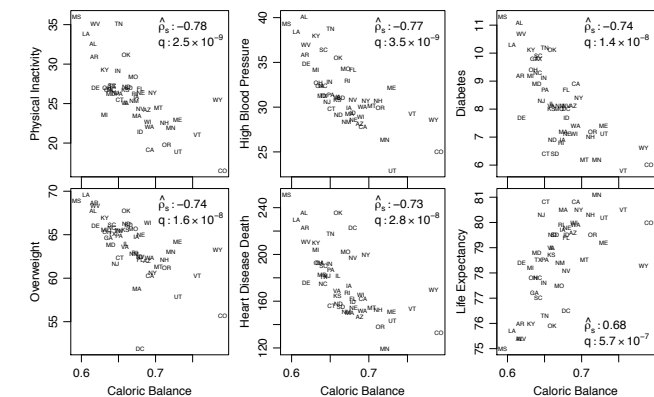# The Lexicocalorimeter:

TABLE S1. Identical to Tab. 1 but with liquids included. Spearman correlation coefficients, $\hat{\rho}_s$, and Benjamini-Hochberg q-values for caloric input C_{in}, caloric output C_{out}, and caloric ratio $C_{rat} = C_{out}/C_{in}$ and demographic data related to food and physical activity, Big Five personality traits [31], health and well-being rankings by state, and socioeconomic status, correlated, ordered from strongest to weakest Spearman correlations with caloric ratio.

The PoCSverse
Lexicocalorimeter
23 of 27

The LCM

Other lexical
meters

References

The Lexicocalorimeter:

TABLE S3. Identical to Tab. 1 but including liquids and using a caloric difference rather than caloric ratio. Spearman correlation coefficients, $\hat{\rho}_s$, and Benjamini-Hochberg q-values for caloric input C_{in}, caloric output C_{out}, and caloric difference $C_{diff}(\alpha) = \alpha C_{out} - (1 - \alpha)C_{in}$ and demographic data related to food and physical activity, Big Five personality traits [31], health and well-being rankings by state, and socioeconomic status, correlated, ordered from strongest to weakest Spearman correlations with caloric ratio. We chose α so that the average of C_{out} matched the average of αC_{in}.

The PoCSverse
Lexicocalorimeter
24 of 27

The LCM

Other lexical
meters

References

The Insomniometer:

Percent of 'sleepy' tweets in state

Correlation with CDC's Sleep Insufficiency measure:
$r_s = .48, p = 5.3 \times 10^{-4}$.

The PoCSverse
Lexicocalorimeter
25 of 27

The LCM

Other lexical
meters

References

The Hangoverometer:

Percent of 'hungover' tweets in state

Correlation with Binge drinking estimates: $r_s = .72, p = 5.8 \times 10^{-9}$.

The PoCSverse
Lexicocalorimeter
26 of 27

The LCM

Other lexical
meters

References

References I

[1] S. E. Alajajian, J. R. Williams, A. J. Reagan, S. C. Alajajian, M. R. Frank, L. Mitchell, J. Lahne, C. M. Danforth, and P. S. Dodds.
The Lexicocalorimeter: Gauging public health through caloric input and output on social media.
PLoS ONE, 12:e0168893, 2017.
arXiv version available at https://arxiv.org/abs/1507.05098. pdf

The PoCSverse
Lexicocalorimeter
27 of 27

The LCM

Other lexical
meters

References