

Lexicocalorimeter: Real-time health measurement

Last updated: 2021/10/26, 20:17:13 EDT

Principles of Complex Systems, Vols. 1 & 2
CSYS/MATH 300 and 303, 2021–2022 | @pocsvox

Prof. Peter Sheridan Dodds | @peterdodds

Computational Story Lab | Vermont Complex Systems Center
Vermont Advanced Computing Core | University of Vermont

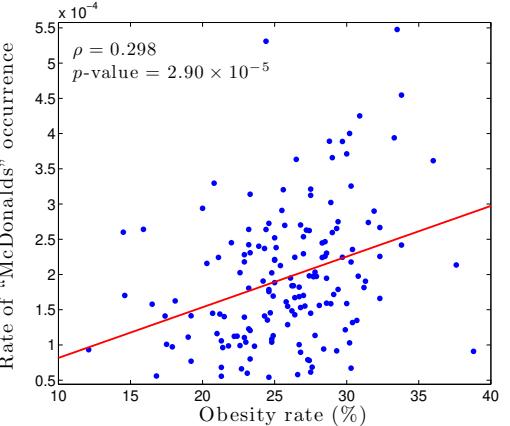


Licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License.

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References

Obesity and tweets—“McDonalds”:



1 of 27

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References

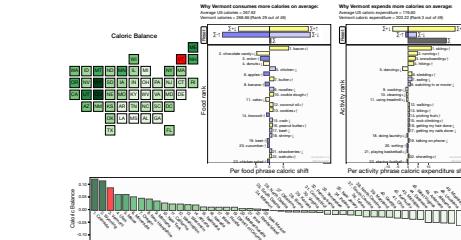
“The Lexicocalorimeter: Gauging public health through caloric input and output on social media”
Alajajian, Williams, Reagan, Alajajian, Frank, Mitchell, Lahne, Danforth, and Dodds. PLoS ONE, 12, e0168893, 2017.



Alajajian, Williams, Reagan, Alajajian, Frank, Mitchell, Lahne, Danforth, and Dodds. PLoS ONE, 12, e0168893, 2017.

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References



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

2 of 27

Outline

The LCM

Other lexical meters

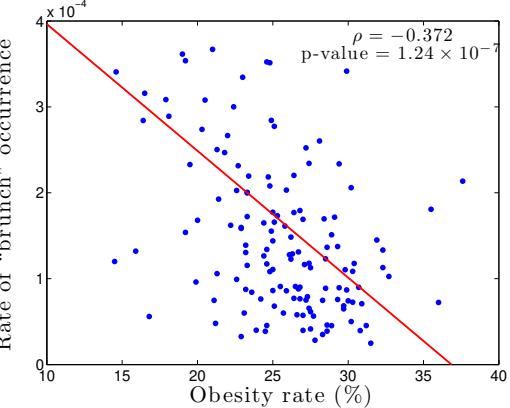
References

2 of 27

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References

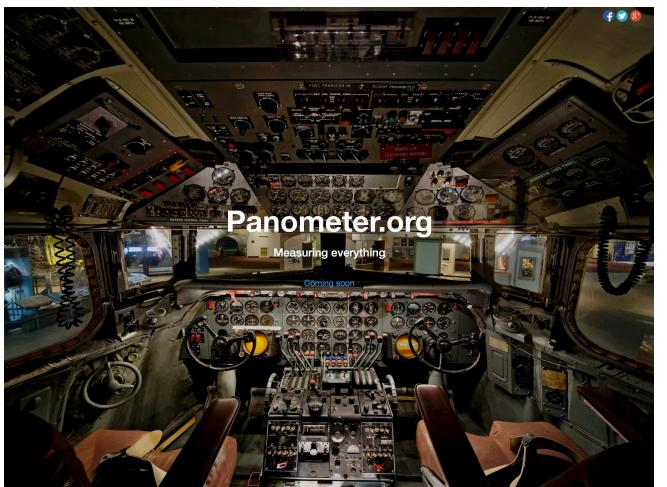
Obesity and tweets—“Brunch”:



2 of 27

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References



Word	ρ	$p\text{-value}$
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}

6 of 27

PoCS
@pocsvox
Lexicocalorimeter

The LCM
Other lexical meters
References

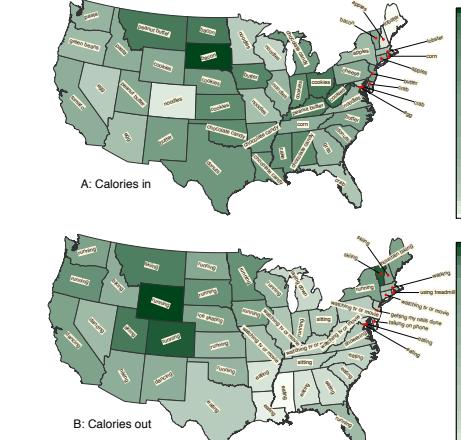
“The Lexicocalorimeter: Gauging public health through caloric input and output on social media”
Alajajian, Williams, Reagan, Alajajian, Frank, Mitchell, Lahne, Danforth, and Dodds. PLoS ONE, 12, e0168893, 2017.



Alajajian, Williams, Reagan, Alajajian, Frank, Mitchell, Lahne, Danforth, and Dodds. PLoS ONE, 12, e0168893, 2017.

PoCS
@pocsvox
Lexicocalorimeter

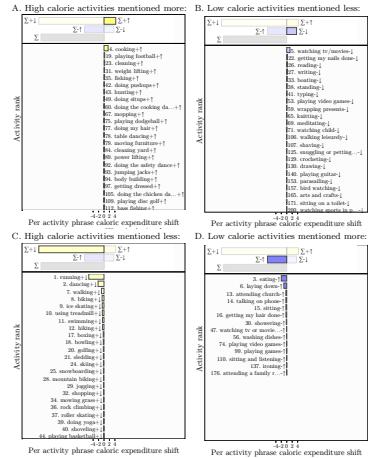
The LCM
Other lexical meters
References



7 of 27

The Lexicocalorimeter:

Four views of activity phrase shifts for Mississippi



PoCS
@pocsvox
Lexicocalorimeter

The Lexicocalorimeter:

The LCM
Other lexical meters

References

Health and/or well-being quantity $\rho_{\text{cal}} \text{ for } C_{\text{cal}}$ $\rho_{\text{cal}} \text{ for } C_{\text{dem}}$ $\rho_{\text{cal}} \text{ for } C_{\text{act}}$ $\rho_{\text{cal}} \text{ for } C_{\text{soc}}$ $\rho_{\text{cal}} \text{ for } C_{\text{env}}$

Health and/or well-being quantity	$\rho_{\text{cal}} \text{ for } C_{\text{cal}}$	$\rho_{\text{cal}} \text{ for } C_{\text{dem}}$	$\rho_{\text{cal}} \text{ for } C_{\text{act}}$	$\rho_{\text{cal}} \text{ for } C_{\text{soc}}$	$\rho_{\text{cal}} \text{ for } C_{\text{env}}$
1. % no physical activity in past 30 days [24]	-0.78	-0.67	-0.58	-0.51	-0.49
2. % have physically active in past 30 days [24]	-0.78	-0.67	-0.58	-0.51	-0.49
3. % high blood pressure [24]	-0.77	-0.69	-0.59	-0.52	-0.49
4. % heart disease death rate [27]	-0.77	-0.69	-0.59	-0.52	-0.49
5. CDC's quality of life ranking [26]	-0.75	-0.69	-0.57	-0.50	-0.48
6. % child overweight/obesity [27]	-0.75	-0.69	-0.57	-0.50	-0.48
7. % adult diabetes rate [27]	-0.75	-0.69	-0.57	-0.50	-0.48
8. % smoking rate [27]	-0.75	-0.69	-0.57	-0.50	-0.48
9. % median household income [27]	-0.75	-0.69	-0.57	-0.50	-0.48
10. % food availability [26]	-0.75	-0.69	-0.57	-0.50	-0.48
11. % total cholesterol [27]	-0.75	-0.69	-0.57	-0.50	-0.48
12. % who eat less than once a day [28]	-0.72	-0.60	-0.61	-0.53	-0.50
13. % child overweight/obesity [27]	-0.72	-0.60	-0.61	-0.53	-0.50
14. % who eat more than once a day [28]	-0.68	-0.53	-0.54	-0.46	-0.43
15. Median daily intake of fruits [28]	-0.65	-0.48	-0.49	-0.41	-0.38
16. Smoking rate [27]	-0.65	-0.48	-0.49	-0.41	-0.38
17. Median household income [27]	-0.65	-0.48	-0.49	-0.41	-0.38
18. % food availability [26]	-0.65	-0.48	-0.49	-0.41	-0.38
19. % total cholesterol [27]	-0.65	-0.48	-0.49	-0.41	-0.38
20. Brain health ranking [28] (lower is better)	-0.65	-0.48	-0.49	-0.41	-0.38
21. % who eat less than once a day [28]	-0.64	-0.47	-0.48	-0.40	-0.37
22. Cigarette consumer rate [27]	-0.64	-0.47	-0.48	-0.40	-0.37
23. US Crime Gun index score [30] (lower is better)	-0.64	-0.47	-0.48	-0.40	-0.37
24. Avg # poor physical health days, past 30 days [24]	-0.62	-0.45	-0.46	-0.38	-0.35
25. Extraversion Big Five personality trait [31]	-0.62	-0.45	-0.46	-0.38	-0.35
26. Binge drinking rate [26]	-0.57	-0.40	-0.41	-0.33	-0.30
27. Avg # poor physical health days, past 30 days [24]	-0.57	-0.40	-0.41	-0.33	-0.30
28. Smoking rate [27]	-0.57	-0.40	-0.41	-0.33	-0.30
29. Smoking rate [27]	-0.57	-0.40	-0.41	-0.33	-0.30
30. Smoking rate [27]	-0.57	-0.40	-0.41	-0.33	-0.30
31. % schools offering fruit/veg at celebrations [28]	-0.56	-0.40	-0.41	-0.33	-0.30
32. % schools offering fruit/veg at celebrations [28]	-0.56	-0.40	-0.41	-0.33	-0.30
33. % overweight/obesity [27]	-0.56	-0.40	-0.41	-0.33	-0.30
34. Conscientiousness Big Five personality trait [31]	-0.56	-0.40	-0.41	-0.33	-0.30
35. % food availability [26]	-0.56	-0.40	-0.41	-0.33	-0.30
36. George Mason overall freedom ranking [32] (lower is free)	-0.56	-0.40	-0.41	-0.33	-0.30
37. Agreeableness Big Five personality trait [31]	-0.56	-0.40	-0.41	-0.33	-0.30

TABLE S1. Identical to Tab. I but with liquids included. Spearman correlation coefficients, ρ_{cal} , and Benjamin-Hochberg q -values for caloric input C_{cal} , caloric output C_{act} , 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 from strongest to weakest Spearman correlations with caloric ratio.

The Lexicocalorimeter:

Health and/or well-being quantity

Health and/or well-being quantity	$\rho_{\text{cal}} \text{ for } C_{\text{cal}}$	$\rho_{\text{cal}} \text{ for } C_{\text{dem}}$	$\rho_{\text{cal}} \text{ for } C_{\text{act}}$	$\rho_{\text{cal}} \text{ for } C_{\text{soc}}$	$\rho_{\text{cal}} \text{ for } C_{\text{env}}$
1. % no physical activity in past 30 days [24]	-0.73	-0.49	-0.63	-0.54	-0.51
2. % have physically active in past 30 days [24]	-0.73	-0.49	-0.63	-0.54	-0.51
3. % high blood pressure [24]	-0.72	-0.48	-0.62	-0.53	-0.50
4. % heart disease death rate [27]	-0.72	-0.48	-0.62	-0.53	-0.50
5. CDC's quality of life ranking [26]	-0.72	-0.48	-0.62	-0.53	-0.50
6. % child overweight/obesity [27]	-0.72	-0.48	-0.62	-0.53	-0.50
7. % adult diabetes rate [27]	-0.72	-0.48	-0.62	-0.53	-0.50
8. % smoking rate [27]	-0.72	-0.48	-0.62	-0.53	-0.50
9. % median household income [27]	-0.72	-0.48	-0.62	-0.53	-0.50
10. % food availability [26]	-0.72	-0.48	-0.62	-0.53	-0.50
11. % total cholesterol [27]	-0.72	-0.48	-0.62	-0.53	-0.50
12. % who eat less than once a day [28]	-0.71	-0.47	-0.61	-0.52	-0.49
13. % child overweight/obesity [27]	-0.71	-0.47	-0.61	-0.52	-0.49
14. % who eat more than once a day [28]	-0.68	-0.43	-0.54	-0.45	-0.42
15. Median daily intake of fruits [28]	-0.65	-0.40	-0.51	-0.42	-0.39
16. Smoking rate [27]	-0.65	-0.40	-0.51	-0.42	-0.39
17. Median household income [27]	-0.65	-0.40	-0.51	-0.42	-0.39
18. % food availability [26]	-0.65	-0.40	-0.51	-0.42	-0.39
19. % total cholesterol [27]	-0.65	-0.40	-0.51	-0.42	-0.39
20. Brain health ranking [28] (lower is better)	-0.65	-0.40	-0.51	-0.42	-0.39
21. % who eat less than once a day [28]	-0.64	-0.40	-0.51	-0.42	-0.39
22. Cigarette consumer rate [27]	-0.64	-0.40	-0.51	-0.42	-0.39
23. US Crime Gun index score [30] (lower is better)	-0.64	-0.40	-0.51	-0.42	-0.39
24. Avg # poor physical health days, past 30 days [24]	-0.62	-0.37	-0.42	-0.33	-0.30
25. Extraversion Big Five personality trait [31]	-0.62	-0.37	-0.42	-0.33	-0.30
26. Binge drinking rate [26]	-0.57	-0.36	-0.41	-0.32	-0.29
27. Avg # poor physical health days, past 30 days [24]	-0.57	-0.36	-0.41	-0.32	-0.29
28. Smoking rate [27]	-0.57	-0.36	-0.41	-0.32	-0.29
29. Smoking rate [27]	-0.57	-0.36	-0.41	-0.32	-0.29
30. Smoking rate [27]	-0.57	-0.36	-0.41	-0.32	-0.29
31. % schools offering fruit/veg at celebrations [28]	-0.56	-0.36	-0.41	-0.32	-0.29
32. % schools offering fruit/veg at celebrations [28]	-0.56	-0.36	-0.41	-0.32	-0.29
33. % overweight/obesity [27]	-0.56	-0.36	-0.41	-0.32	-0.29
34. Conscientiousness Big Five personality trait [31]	-0.56	-0.36	-0.41	-0.32	-0.29
35. % food availability [26]	-0.56	-0.36	-0.41	-0.32	-0.29
36. George Mason overall freedom ranking [32] (lower is free)	-0.56	-0.36	-0.41	-0.32	-0.29
37. Agreeableness Big Five personality trait [31]	-0.56	-0.36	-0.41	-0.32	-0.29

TABLE S1. Spearman correlation coefficients, ρ_{cal} , and Benjamin-Hochberg q -values for caloric input C_{cal} , caloric output C_{act} , 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 from strongest to weakest Spearman correlations with caloric ratio. The bottom 9 quantities were not significantly correlated with C_{cal} according to our test. See Table S1, S2, and S3 present the same analysis for caloric measures including phrases representing liquids, and the difference $C_{\text{cal}}(n) = C_{\text{cal}}(n) - (1 - n)C_{\text{act}}$, both without and with liquids included.

PoCS
@pocsvox
Lexicocalorimeter

The Lexicocalorimeter:

The LCM
Other lexical meters

References

Health and/or well-being quantity

Health and/or well-being quantity	$\rho_{\text{cal}} \text{ for } C_{\text{cal}}$	$\rho_{\text{cal}} \text{ for } C_{\text{dem}}$	$\rho_{\text{cal}} \text{ for } C_{\text{act}}$	$\rho_{\text{cal}} \text{ for } C_{\text{soc}}$	$\rho_{\text{cal}} \text{ for } C_{\text{env}}$
1. % no physical activity in past 30 days [24]	-0.73	-0.42	-0.63	-0.54	-0.51
2. % have physically active in past 30 days [24]	-0.73	-0.42	-0.63	-0.54	-0.51
3. % high blood pressure [24]	-0.72	-0.40	-0.62	-0.53	-0.50
4. % heart disease death rate [27]	-0.72	-0.40	-0.62	-0.53	-0.50
5. CDC's quality of life ranking [26]	-0.72	-0.40	-0.62	-0.53	-0.50
6. % child overweight/obesity [27]	-0.72	-0.40	-0.62	-0.53	-0.50
7. % adult diabetes rate [27]	-0.72	-0.40	-0.62	-0.53	-0.50
8. % smoking rate [27]	-0.72	-0.40	-0.62	-0.53	-0.50
9. % median household income [27]	-0.72	-0.40	-0.62	-0.53	-0.50
10. % food availability [26]	-0.72	-0.40	-0.62	-0.53	-0.50
11. % total cholesterol [27]	-0.72	-0.40	-0.62	-0.53	-0.50
12. % who eat less than once a day [28]	-0.71	-0.39	-0.61	-0.52	-0.49
13. % child overweight/obesity [27]	-0.71	-0.39	-0.61	-0.52	-0.49
14. % who eat more than once a day [28]	-0.68	-0.36	-0.58	-0.49	-0.46
15. Median daily intake of fruits [28]	-0.65	-0.34	-0.55	-0.47	-0.44
16. Smoking rate [27]	-0.65	-0.34	-0.55	-0.47	-0.44
17. Median household income [27]	-0.65	-0.34	-0.55	-0.47	-0.44
18. % food availability [26]	-0.65	-0.34	-0.55	-0.47	-0.44
19. % total cholesterol [27]	-0.65	-0.34	-0.55	-0.47	-0.44
20. Brain health ranking [28] (lower is better)	-0.65	-0.34	-0.55	-0.47	-0.44
21. % who eat less than once a day [28]	-0.64	-0.34	-0.55	-0.47	-0.44
22. Cigarette consumer rate [27]	-0.64	-0.34	-0.55	-0.47	-0.44
23. US Crime Gun index score [30] (lower is better)	-0.64	-0.34	-0.55	-0.47	-0.44
24. Avg # poor physical health days, past 30 days [24]	-0.62	-0.33	-0.54	-0.46	-0.43
25. Extraversion Big Five personality trait [31]	-0.62	-0.33	-0.54	-0.46	-0.43
26. Binge drinking rate [26]	-0.57	-0.30	-0.51	-0.43	-0.40
27. Avg # poor physical health days, past 30 days [24]	-0.57	-0.30	-0.51	-0.43	-0.40
28. Smoking rate [27]	-0.57	-0.30	-0.51	-0.43	-0.40
29. Smoking rate [27]	-0.57	-0.30	-0.51	-0.43	-0.40
30. Smoking rate [27]	-0.57	-0.30	-0.51	-0.43	-0.40
31. % schools offering fruit/veg at celebrations [28]	-0.56	-0.30	-0.51	-0.43	-0.40
32. % schools offering fruit/veg at celebrations [28]	-0.56	-0.30	-0.51	-0.43	-0.40
33. % overweight/obesity [27]	-0.56	-0.30	-0.51	-0.43	-0.40
34. Conscientiousness Big Five personality trait [31]	-0.56	-0.30	-0.51	-0.43	-0.40
35. % food availability [26]	-0.56	-0.30	-0.51	-0.43	-0.40
36. George Mason overall freedom ranking [32] (lower is free)	-0.56	-0.30	-0.51	-0.43	-0.40
37. Agreeableness Big Five personality trait [31]	-0.56	-0.30	-0.51	-0.43	-0.40

TABLE S3. Identical to Tab. I but including liquids and using a caloric difference rather than caloric ratio. Spearman correlation coefficients, ρ_{cal} , and Benjamin-Hochberg q -values for caloric input C_{cal} , caloric output C_{act} , 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 from strongest to weakest Spearman correlations with caloric ratio. We chose so that the average of C_{act} matched the average of C_{cal} .

PoCS
@pocsvox
Lexicocalorimeter

References

21. % schools offering fruit/veg at celebrations [28]

22. % schools offering fruit/veg at celebrations [28]

23. % overweight/obesity [27]

24. Conscientiousness Big Five personality trait [31]

25. % food availability [26]

26. George Mason overall freedom ranking [32] (lower is free)

27. Agreeableness Big Five personality trait [31]

28. George Mason overall freedom ranking [32] (lower is free)

29. Conscientiousness Big Five personality trait [31]

30. % overweight/obesity [27]

31. % food availability [26]

32. % schools offering fruit/veg at celebrations [28]

33. % overweight/obesity [27]

34. Conscientiousness Big Five personality trait [31]

35. % food availability [26]

36. George Mason overall freedom ranking [32] (lower is free)

37. Agreeableness Big Five personality trait [31]

38. Agreeableness Big Five personality trait [31]

39. Conscientiousness Big Five personality trait [31]