Pathos: Associations of Word Usage and Emotions

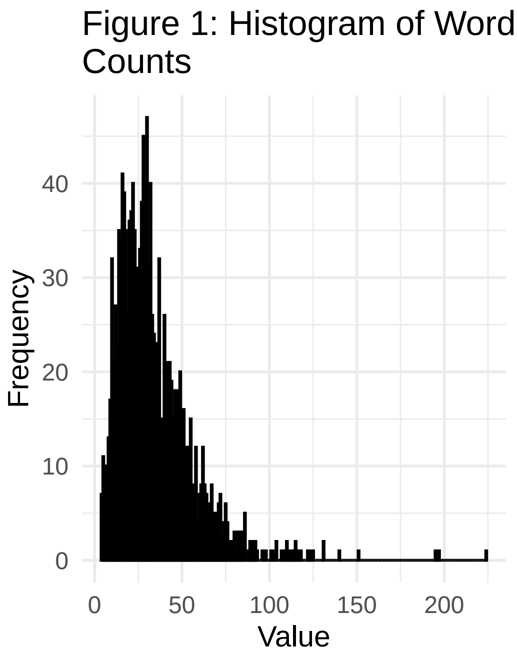
Max Melnikas

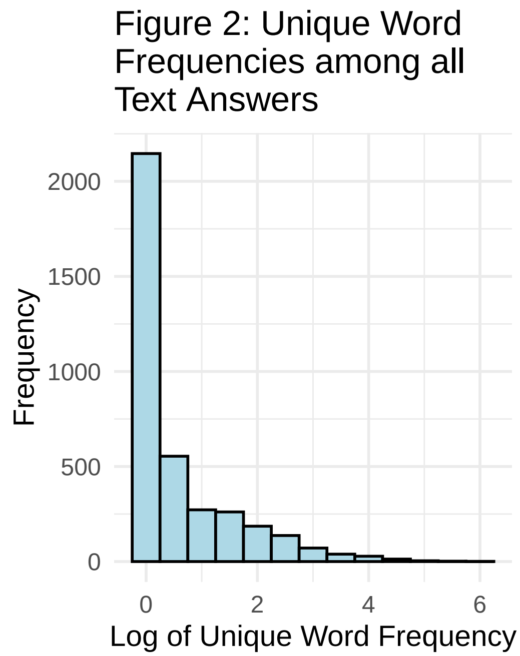
Introduction

In the realm of psychology and neuroscience, understanding human experiences and emotions through word usage can present a fascinating and difficult challenge. Words choice can be highly person and context-dependent. However, with a large enough sample of written answers to a single question prompt, we may be able to identify certain trends in word usage.

In this project, I used a dataset (X. Alice Li and Devi Parikh, 2019) that contains a large number (N = 1473) of written responses to the question: “What were salient aspects of your day yesterday? How did you feel about them?”. Additionally, each response is labelled with one or more emotion from an exhaustive list of 18 different emotions. Below are two figures that were created in the exploratory data analysis stage. Figure 1 depicts a histogram of word counts from the text responses while Figure 2 shows a log-transformed histogram of unique word frequencies from the entire pool of responses. As we can see from Figure 2, the tall bar at zero on the x-axis indicates that a majority of words were only observed once in the entire dataset. On the other hand, some words were observed very frequently, exemplified by the long tail of the histogram.

In my analysis, I will attempt to find associations between frequent words that participants included in their responses and the emotions these responses were labelled with. Next, I will shift my analysis to an exploration of word co-occurence and whether certain word pairs have stronger associations to emotions than the single words on their own. Overall, I hope to answer the question of whether word usage can inform emotional state.





Methods

To commence my analysis, I had to conduct a series of preprocessing steps. Each entry (row) of the initial dataset included a response to the question as a string, followed by 18 indicator variables - one for each listed emotion. Firstly, I had to seperate each string answer into word tokens and account for capitalization and punctuation. Additionally, I filtered out certain uninformative words, commonly referred to as “stop words”, out of my data dictionary. Stop words are words such as “the” or “and” that occur very frequently in text but are not reasonably associated with any emotions. Furthermore, from Figure 1 we can see that certain responses had low word counts. After filtering out stop words, a handful of responses were no longer viable for exploring word co-occurrence. For this reason, observations 877 and 1448 were removed from the dataset leaving a final sample size of 1471 observations.

The intended outcome of this study were labelled emotions self-reported by the user from a list of 18 options. Working with 18 different categories was a challenge because of the variance of the category counts. Table 1 below contains these counts, the range of which extended from jealous (n = 3) to happy (n = 730). Therefore, I attempted to reduce the dimensionality of these 18 outcomes. Initially I attempted a dichotomization of the 18 categories into just two partitions: positive and negative. I found this approach to be not sufficient as these groupings of emotions were overly crude and failed to distinguish between significant differences. Next, I contemplated using a reductional technique such as PCA. The first nine principle components were able to account for 90% percent of the variance. However completing PCA on my outcomes would complicate the interpretations of my results, therefore I decided not to

Table 1: Initial Catgeory Counts

	.
Afraid	18
Angry	28
Anxious	125
Ashamed	17
Awkward	15
Bored	49
Calm	368
Confused	28
Disgusted	22
Excited	250
Frustrated	140
Happy	730
Jealous	3
Nostalgic	61
Proud	337
Sad	43
Satisfied	591
Surprised	64

go with this approach. I finally settled on an approach rooted in psychology domain knowledge: I used the Wilcox Feeling Wheel to aggregate my outcomes into six final categories: mad, sad, joyful, scared, peaceful and powerful (Gloria Wilcox, 1982). This transformation is depicted in Table 2 below. All new categories have more than 100 observations in them, making associational techniques much more applicable.

To explore associations between word usage and emotion labels, I will regress single word and word pair occurrence onto the indicator variables for each emotion using logistic regression. To handle the computational complexity of this task, I will identify the most frequently occurring words and word pairs in this dataset to use as predictors. Next I will use Lasso, a penalized

Table 2: Transformed Catgeory Counts

	Old.Categories	Count
sad	Ashamed, Bored, Sad	101
mad	Angry, Disgusted, Frustrated, Jealous	166
joyful	Excited, Happy	808
scared	Afraid, Anxious, Awkward, Confused	158
peaceful	Calm, Nostalgic, Satisfied	830
powerful	Proud, Surprised	391

regression technique, to identify the most informative associations (Robert Tibshirani, 1996). In doing so, I will be able to estimate odds ratios of experiencing emotions when using certain words or pairs.

To find the most frequent words I simply browsed for the 95% quantile of most frequently occurring words which turned out to be ones that appeared more than 14 times. To find frequent word pairs, I employed the Apriori algorithm which is a classical associational mining technique that utilizes efficient subsetting (Rakesh Agrawal and Ramakrishnan Srikant, 1994). The first step in this approach is to identify the most frequently occurring items at some threshold (in this instance being 14). Then all possible combinations from those most frequent candidates are identified and sampled. After that, any word pairs that are below that same threshold are eliminated, and the resulting word pairs will be the most frequent in the dataset. At the current threshold, 63 final word pairs were identified with "time" and "happy" being the most frequent one at a count of 47.

Results

As an initial model attempt, I used all frequent word candidates as predictors in my logistic model. Table 3 below contains the output for this kind of model applied to the Joyful emotion. Additionally it was filtered to only show words that had p-values under 0.05. As we can see in the table, many words satisfy the statistical significance threshold, therefore penalized regression in the form of Lasso was applied to dampen the number of predictors. Additionally, this helps to remove words that are found to be non significant. It appears that using Lasso regression heavily decreases the overall slopes of the words. By exponentiating these slopes, we can determine odds ratios. For example, the slope for the predictor "frustrating" was -0.62 which produces an odds ratio of 0.535. That is to say, the group of people that mentioned the word "frustrating" in their response had 0.535 the odds of being in outcome category "joyful" compared to their counterparts who did not mention the word "frustrating", on average holding all other covariates constant.

Overall, there was a lot of similarity in which predictors lasso determined to be significant for each emotion. The full tables for each emotion are listed in the appendix. To summarize some of the most salient findings: love, fun and excitement are all predictive of joyfulness. Interestingly, it appears as though the word "sleep" had an odds ratio less than one. One explanation for this is that perhaps "sleep" had a strong negative association with the joyful emotion because it was commonly referred to in the context of a lack of sleep. Pain, sickness and tiredness are all predictive of the "mad" state. The "sad" state had strong associations with mentions of money and pay, while the "scared" state had negative associations with mentions of god.

Next, I continued the same analysis using word pairs as predictors rather than just single words. When running simple logistic models, one common finding across all emotions was that far fewer predictors were deemed to be statistically significant at a p-value threshold of

0.05. This may be explained by the lower overall incidence of word pairs as compared to individual words. The joyful category was the most popular outcome of this dataset, and subsequently this emotion had the highest number of significant predictors. The word pair composed of “happy” and “family” had a slope estimate of 1.54 for this emotion, which can be converted to an odds ratio of 4.66. That is to say, the group of people that mentioned the words “happy” and “family” in their response had 4.66 the odds of being in outcome category “joyful” compared to their counterparts who did not mention both of those words, on average holding all other covariates constant.

Other emotions that had fewer counts in the dataset than the joyful emotion benefited greatly from the application of Lasso regression. The mad and sad emotions, for example, only had a single word pair predictor that was found to be significant at the 0.05 significance level. However after fitting a Lasso regression and tuning the lambda parameter, more relevant predictors were identified and interpreted. The joyful emotion was characterized by word pairs such as “time/friends” and “time/love” which were both predictive of a joyful emotional state. One particular word pair that stood out in the Lasso regression model for the mad emotion was “time/didn’t”, indicating that not finding the time to complete various tasks may contribute to a mad emotional state. A notable word pair positive predictor for the peaceful emotion was “time/bed”, further highlighting the close relationship between daily behaviors and emotional states. The sad, scared and powerful outcomes all had low numbers of respondents, causing those models to be less uninformative.

Conclusion

In this project, an efficient subsetting algorithm was applied to text-based observations to find common word pairings. These pairings, along with frequent individual words were then used as predictors in simple logistic and penalized logistic models to draw associations between word usage and emotions. Though originally, the dataset used in this project seemed to have a sufficient sample size (N=1471), the diversity in word choice has highlighted the need for extremely large datasets to conduct reproducible and generalizable word usage studies. Moreover, the entries of this dataset were produced by 500 total participants, indicating that each participant could submit up to three answers. The analysis completed in this project assumes independence in observations, suggesting that the variance may be underestimated resulting in a downward bias in p-values. To adequately adjust for this effect, a clustering appraoch should be considered in future studies. Finally, one last important point of consideration is the selection of stop words. The word “time”, for example, emerges exceedingly frequently in significant word pairs across all emotions. However, on its own, it is often found to be not significant. This finding may suggest that the inclusion of the word “time” in various word pairs in different models may be driven by its frequency in the text answers rather than its association to the emotion.

The analysis of word usage may present many opportunities in behavioral research and the field of public health. In this project, various words and word pairs were associated to several classes of emotions. However, one could foresee a similar study attempting to relate word usage to cognitive functionality. Such an approach may prove useful in mental health monitoring and early detection of cognitive disorders. Overall, studying large datasets of words and text may create unique opportunities to qualify health states and help spark novel public health interventions.

References

Agrawal, R. and Srikant, R. (1994) Fast Algorithms for Mining Association Rules in Large Databases. Proceedings of the 20th International Conference on Very Large Data Bases, VLDB, Santiago de Chile, 12-15 September 1994, 487-499.

Li, A. and Parikh, D. (2020) Lemotif: An Affective Visual Journal Using Deep Neural Networks. arXiv:1903.07766.

Tibshirani, R. (1996), Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological), 58: 267-288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Willcox, G. (1982). The Feeling Wheel A Tool for Expanding Awareness of Emotions and Increasing Spontaneity and Intimacy. Transactional Analysis Journal, 12(4), 274-276.

Appendix

Table 3: Joyful Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-0.5189731	0.1337568	-3.879976	0.0001045
anxious	-2.4006255	0.6761641	-3.550359	0.0003847
calm	-1.9933305	0.5180364	-3.847858	0.0001192
chicken	1.6572496	0.7714049	2.148352	0.0316858
didnt	-1.0170101	0.4547791	-2.236273	0.0253339
dog	1.3906254	0.6959440	1.998186	0.0456965
dont	-1.2764783	0.4789562	-2.665125	0.0076960
excited	4.5563678	1.1971199	3.806108	0.0001412
friends	0.8517320	0.3534474	2.409784	0.0159620
frustrated	-1.8121191	0.6387028	-2.837187	0.0045513
frustrating	-2.5927290	1.0785381	-2.403929	0.0162199
happy	1.8731123	0.2714381	6.900699	0.0000000
life	0.7788886	0.3494747	2.228741	0.0258312
love	1.2304024	0.3237852	3.800057	0.0001447
loved	2.5374749	1.2563128	2.019779	0.0434063
mind	2.1647006	0.8872836	2.439694	0.0146997
morning	0.9703494	0.4346296	2.232590	0.0255760
nice	1.0182112	0.3420408	2.976871	0.0029121
sad	-2.0517590	0.8257617	-2.484687	0.0129666
sense	-1.8153015	0.9082235	-1.998739	0.0456366
set	1.9888678	0.8441012	2.356196	0.0184632
sick	-1.5397438	0.6730403	-2.287744	0.0221524
sleep	-0.7042209	0.3518040	-2.001742	0.0453125
spent	0.9573636	0.4009769	2.387578	0.0169598
times	1.9389224	0.6302121	3.076619	0.0020936
week	0.9288889	0.3926297	2.365814	0.0179905
weeks	1.6057550	0.6803696	2.360122	0.0182689
wife	0.8141615	0.3917735	2.078143	0.0376962

Table 4: Joyful Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.045)

Word	Slope	Odds Ratio
anxious	-0.3696120	0.6910024
calm	-0.0934286	0.9108030
didnt	-0.0222855	0.9779610
dinner	0.0856609	1.0894369
excited	0.6289378	1.8756172
frustrated	-0.0141425	0.9859570
frustrating	-0.6241183	0.5357336
fun	0.1852811	1.2035568
happy	0.7913147	2.2062952
love	0.1398840	1.1501404
sleep	-0.3140473	0.7304845
spent	0.0273026	1.0276788
time	0.0645999	1.0667322
wife	0.0165008	1.0166377

Table 5: Joyful Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
time_spentTRUE	1.422060	0.5115856	2.779711	0.0054407
happy_familyTRUE	1.535577	0.6718407	2.285627	0.0222761
happy_feelTRUE	2.299503	0.7658992	3.002357	0.0026790
time_lotTRUE	-1.217161	0.5511842	-2.208264	0.0272258
time_niceTRUE	1.720142	0.6623285	2.597112	0.0094011
time_friendsTRUE	1.699745	0.6773307	2.509475	0.0120911
time_loveTRUE	1.850018	0.7206803	2.567043	0.0102570
feel_calmTRUE	-1.917067	0.7430955	-2.579839	0.0098846
satisfied_feelTRUE	-1.607668	0.7115996	-2.259232	0.0238690

Table 6: Joyful Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.03)

Word	Slope	Odds Ratio
time_happy	0.3540312	1.4247996
time_spent	0.4413985	1.5548802
time_spend	0.0989050	1.1039615
happy_family	0.5837373	1.7927260
happy_feel	0.5206228	1.6830756
sleep_night	-0.3651781	0.6940730
time_nice	0.2420454	1.2738521
time_dinner	0.1772751	1.1939594
time_friends	0.2180869	1.2436951
time_love	0.1156083	1.1225561
feel_calm	-0.1835003	0.8323516
sleep_nights	-0.0922804	0.9118495
happy_food	0.0954185	1.1001192

Table 7: Mad Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-2.326440	0.2864088	-8.122798	0.0000000
bad	5.034343	1.2906637	3.900584	0.0000960
couldnt	2.427735	1.2029851	2.018093	0.0435816
difficult	3.282446	1.2259038	2.677573	0.0074158
doesnt	2.389306	1.0583763	2.257520	0.0239756
dont	2.106547	0.7463708	2.822387	0.0047668
feel	-2.485119	0.8563402	-2.902023	0.0037076
frustrated	9.350275	1.8109012	5.163327	0.0000002
frustrating	6.662351	1.4926511	4.463435	0.0000081
happy	-2.815978	1.0171232	-2.768571	0.0056303
helped	-7.101505	2.5251953	-2.812260	0.0049195
home	-1.983763	0.8667935	-2.288622	0.0221013
hours	1.841177	0.7557518	2.436219	0.0148417
kids	3.015425	1.1538597	2.613338	0.0089663
pain	3.172505	1.1255070	2.818734	0.0048213
sick	2.939203	0.9941115	2.956613	0.0031104

Table 8: Mad Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.018)

Word	Slope	Odds Ratio
10	0.1678005	1.1827006
bad	0.8551806	2.3517991
couldnt	0.5055038	1.6578204
difficult	0.8853465	2.4238241
doesnt	0.6176846	1.8546288
dont	0.4606403	1.5850886
feel	-0.0038248	0.9961825
frustrated	3.4199196	30.5669563
frustrating	3.7050971	40.6539930
happy	-0.4098015	0.6637820
hour	0.3118930	1.3660086
hours	0.3182911	1.3747765
ill	0.8221205	2.2753195
nice	-0.1827943	0.8329395
pain	1.1283365	3.0905113
pay	0.3377083	1.4017316
sick	1.1048742	3.0188447
tired	0.6657885	1.9460243

Table 9: Mad Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-1.914386	0.0945170	-20.254405	0.0000000
time_lotTRUE	1.579174	0.6258811	2.523122	0.0116318

Table 10: Mad Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.0135)

Word	Slope	Odds Ratio
time_spend	-0.0711134	0.9313563
happy_feel	-0.3497906	0.7048357
time_lot	0.3350803	1.3980526
time_nice	-0.2045761	0.8149927
time_spending	-0.1224501	0.8847500
time_dinner	-0.0833865	0.9199955
makes_happy	-0.0131784	0.9869081
love_family	-0.0000892	0.9999108
time_didnt	0.4050966	1.4994473
sleep_nights	0.0801598	1.0834602
time_im	0.1456584	1.1568009

Table 11: Sad Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-3.415700	0.3739005	-9.135318	0.0000000
ate	2.871284	1.1194430	2.564922	0.0103199
boring	4.782604	1.3326218	3.588868	0.0003321
cooking	4.822497	1.9503484	2.472634	0.0134122
didnt	1.832199	0.8755028	2.092740	0.0363724
doesnt	3.433143	1.1877422	2.890478	0.0038466
frustrating	3.179157	0.9448639	3.364672	0.0007663
god	-4.880309	2.1772644	-2.241487	0.0249946
happy	-3.413964	1.4424938	-2.366710	0.0179470
house	2.087817	1.0283413	2.030276	0.0423284
husband	-3.372479	1.5071036	-2.237722	0.0252392
pay	3.705894	1.3853978	2.674967	0.0074737
person	3.246079	1.3016735	2.493774	0.0126393
productive	2.500459	1.2240239	2.042819	0.0410704
relax	3.015462	1.4478579	2.082706	0.0372780
sad	6.823720	1.6702495	4.085449	0.0000440
special	3.426474	1.6111473	2.126730	0.0334426
super	6.562811	1.9415797	3.380140	0.0007245
time	-1.545582	0.7771944	-1.988669	0.0467378
wasnt	2.581938	1.1501131	2.244943	0.0247718
watching	4.111049	1.9655145	2.091589	0.0364753
weight	2.496080	1.2384143	2.015545	0.0438476

Table 12: Sad Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.018)

Word	Slope	Odds Ratio
bad	0.1201357	1.1276499
boring	2.1614209	8.6834676
couldnt	0.9848517	2.6774147
doesnt	0.6517359	1.9188690
dont	0.1156696	1.1226249
fairly	1.4181453	4.1294544
feeling	0.0051514	1.0051647
frustrated	0.4081852	1.5040858
frustrating	1.1697106	3.2210604
happy	-0.1261389	0.8814924
money	0.2780624	1.3205687
pay	0.7433711	2.1030131
sad	2.4303599	11.3629710
sick	0.4162848	1.5163177
super	0.3374999	1.4014395
wasnt	0.0140768	1.0141763
weight	0.1711132	1.1866251

Table 13: Sad Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-2.450430	0.1182973	-20.714167	0.0000000
lot_feelTRUE	1.714255	0.8042376	2.131528	0.0330457

Table 14: Sad Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.008)

Word	Slope	Odds Ratio
time_spent	-0.4189611	0.6577298
happy_family	-0.2370772	0.7889304
makes_feel	-0.0031888	0.9968163
happy_feel	-0.4113393	0.6627620
family_dinner	-0.1432566	0.8665317
time_spending	-0.1450583	0.8649719
makes_happy	-0.0114319	0.9886332
time_friends	-0.2000753	0.8186691
lot_feel	1.0842082	2.9570976
time_didnt	0.2564577	1.2923441
sleep_nights	-0.0248124	0.9754929

Table 15: Scared Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-2.447915	0.2529365	-9.677981	0.0000000
anxious	6.967911	1.1288700	6.172466	0.0000000
difficult	3.266935	1.1777034	2.773988	0.0055374
doesnt	2.346065	0.9790354	2.396302	0.0165614
dont	2.240047	0.6501550	3.445404	0.0005702
exercise	-2.514444	1.1268285	-2.231434	0.0256524
finally	-4.145259	1.8533144	-2.236674	0.0253077
found	1.509279	0.7478251	2.018224	0.0435679
frustrated	2.948761	0.8048071	3.663935	0.0002484
god	-2.939163	1.3487310	-2.179206	0.0293164
happy	-2.286129	0.7219236	-3.166719	0.0015417
job	1.248440	0.5624629	2.219595	0.0264463
makes	-2.268991	0.9740194	-2.329513	0.0198319
meet	2.931863	1.1796490	2.485369	0.0129417
mom	1.864452	0.9403129	1.982800	0.0473898
pretty	1.948708	0.6857755	2.841612	0.0044886
salient	1.919839	0.8735306	2.197792	0.0279639
taking	2.195257	1.0702853	2.051095	0.0402577
weight	2.588116	1.0600059	2.441606	0.0146221
woke	2.923973	1.1707499	2.497521	0.0125065

Table 16: Scared Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.025)

Word	Slope	Odds Ratio
10	0.0256698	1.0260021
anxious	3.1293313	22.8586895
bad	0.0821113	1.0855766
couldnt	0.5532542	1.7389026
difficult	0.0012933	1.0012941
doesnt	0.0150327	1.0151463
dont	0.6811284	1.9761063
found	0.1756457	1.1920157
frustrated	0.9278641	2.5291015
frustrating	0.6271718	1.8723078
happy	-0.1734111	0.8407919
im	0.0012895	1.0012903
job	0.0264173	1.0267693
lot	0.0021667	1.0021691
nice	-0.1027918	0.9023148
normal	0.3568296	1.4287924
pain	0.1754761	1.1918135
slow	0.0297980	1.0302464

Table 17: Scared Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-1.938938	0.0958016	-20.239085	0.0000000
time_familyTRUE	2.363709	1.0963773	2.155927	0.0310894
time_lotTRUE	1.529255	0.6854402	2.231055	0.0256775
im_feelTRUE	2.715467	1.1206903	2.423031	0.0153916
night_hoursTRUE	2.875792	1.3471524	2.134719	0.0327840

Table 18: Scared Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.012)

Word	Slope	Odds Ratio
time_happy	-0.3863435	0.6795371
time_spend	-0.3652063	0.6940535
happy_family	-0.2702013	0.7632259
makes_feel	-0.0303854	0.9700716
time_lot	0.0068076	1.0068309
time_nice	-0.1146566	0.8916723
time_bed	-0.0094288	0.9906155
wife_time	-0.0071194	0.9929059
wife_dinner	-0.0011054	0.9988952
slept_sleep	0.1539088	1.1663845
night_hours	0.3271769	1.3870468

Table 19: Peaceful Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-0.4865905	0.1284441	-3.788345	0.0001517
anxious	-1.8809499	0.5799326	-3.243394	0.0011811
bad	-1.5584449	0.5946638	-2.620716	0.0087745
bought	3.0478459	1.3382799	2.277435	0.0227602
calm	4.6478090	1.4361857	3.236217	0.0012113
catch	2.4195184	0.8329411	2.904789	0.0036750
delicious	0.9592612	0.4552045	2.107319	0.0350899
eat	-0.8889761	0.4176819	-2.128357	0.0333075
enjoyable	2.5513010	0.9417212	2.709189	0.0067448
feel	1.0101045	0.2430909	4.155255	0.0000325
frustrated	-2.2762816	0.8177412	-2.783621	0.0053756
frustrating	-3.8005214	1.1356519	-3.346555	0.0008182
god	1.1417365	0.3789338	3.013024	0.0025866
home	0.6217105	0.2923608	2.126518	0.0334602
house	1.3243784	0.5231279	2.531653	0.0113526
minutes	-2.5055654	0.8151392	-3.073788	0.0021136
relaxing	1.6863769	0.6722902	2.508406	0.0121277
satisfied	4.2370846	1.0866062	3.899375	0.0000964
sleep	1.5072740	0.3758214	4.010612	0.0000606
super	-1.9588320	0.7455809	-2.627256	0.0086076
time	0.4136117	0.2026468	2.041047	0.0412462
times	1.3129543	0.5457615	2.405729	0.0161402
watch	1.8774033	0.8403945	2.233955	0.0254861
watched	1.6280384	0.6897024	2.360494	0.0182506
weight	-1.2967241	0.6253349	-2.073647	0.0381121
wife	0.9674441	0.3478728	2.781028	0.0054187

Table 20: Peaceful Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.03)

Word	Slope	Odds Ratio
anxious	-0.6144996	0.5409115
bought	0.0786662	1.0818431
calm	0.9488141	2.5826452
catch	0.0994663	1.1045812
enjoyable	0.1411504	1.1515978
feel	0.3437302	1.4101980
frustrated	-0.6921661	0.5004908
frustrating	-1.1251593	0.3246008
god	0.0900169	1.0941928
house	0.0087711	1.0088097
nice	0.1192631	1.1266663
night	0.1021957	1.1076002
ready	0.1787504	1.1957222
relax	0.1368942	1.1467069
relaxing	0.4409565	1.5541930
routine	0.0436423	1.0446086
satisfied	1.3041544	3.6845722
sleep	0.3169188	1.3728912
time	0.0254841	1.0258116
times	0.0278004	1.0281905
told	-0.0129695	0.9871143

Table 21: Peaceful Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
makes_feelTRUE	0.943871	0.4640814	2.033848	0.0419669
sleep_nightTRUE	1.341865	0.5302975	2.530401	0.0113932
wife_dinnerTRUE	1.609648	0.7466707	2.155767	0.0311018
time_lifeTRUE	2.270246	1.0060327	2.256632	0.0240311

Table 22: Peaceful Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.023)

Word	Slope	Odds Ratio
happy__family	0.0133057	1.013395
makes__feel	0.3120264	1.366191
happy__feel	0.0537336	1.055203
sleep__night	0.3335506	1.395916
time__bed	0.5345777	1.706727
time__enjoy	0.0308492	1.031330
spend__family	0.0799256	1.083206
wife__dinner	0.1821391	1.199781
feel__calm	0.9418199	2.564645
satisfied__feel	0.9375523	2.553723
home__happy	0.0887589	1.092817
time__life	0.3159106	1.371508

Table 23: Powerful Emotion Logistic Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-1.1682201	0.1428315	-8.179010	0.0000000
anxious	-1.5384111	0.6372421	-2.414171	0.0157711
calm	-1.7121113	0.7350660	-2.329194	0.0198488
eating	-2.6438008	1.2654709	-2.089183	0.0366912
exercising	1.7686823	0.7909202	2.236234	0.0253365
feeling	0.8074990	0.3987914	2.024866	0.0428811
goal	1.2416987	0.6259089	1.983833	0.0472745
past	2.7448942	0.7422901	3.697873	0.0002174
pleased	2.0107283	0.8169220	2.461347	0.0138417
proud	3.3679521	0.4397661	7.658508	0.0000000
sick	-2.0765150	0.9036169	-2.298004	0.0215616
sleep	-1.0430295	0.4758659	-2.191856	0.0283899
surprised	2.8682254	0.7087839	4.046686	0.0000519
workout	0.9601753	0.4847856	1.980619	0.0476341

Table 24: Powerful Emotion Lasso Predictor Slopes and Odds Ratios (Lambda = 0.025)

Word	Slope	Odds Ratio
calm	-0.0803549	0.9227888
couldnt	-0.1960658	0.8219581
exercising	0.2083801	1.2316812
food	-0.1537735	0.8574662
goal	0.0287153	1.0291315
past	0.1575710	1.1706638
pleased	0.5569253	1.7452979
project	0.0472694	1.0484044
proud	2.3140723	10.1155340
sleep	-0.3204693	0.7258083
surprised	1.1035791	3.0149374
workout	0.0293347	1.0297692

Table 25: Powerful Emotion Logistic Word Pair Predictor Slopes and Odds Ratios

term	estimate	std.error	statistic	p.value
(Intercept)	-1.029020	0.0698182	-14.738559	0.0000000
sleep__nightTRUE	-1.752025	0.8917363	-1.964735	0.0494450
time__nightTRUE	-1.613652	0.7360888	-2.192198	0.0283652
lot__feelTRUE	-2.585705	1.2557298	-2.059125	0.0394823
proud__happyTRUE	3.550482	0.9356340	3.794734	0.0001478
life__familyTRUE	1.494260	0.7417761	2.014435	0.0439638
life__happyTRUE	-2.305154	1.1753895	-1.961183	0.0498577
time__proudTRUE	2.844286	0.9017410	3.154216	0.0016093
time__imTRUE	1.430905	0.6647707	2.152479	0.0313596

Table 26: Powerful Emotion Lasso Word Pair Predictor Slopes and Odds Ratios (Lambda = 0.0165)

Word	Slope	Odds Ratio
sleep__night	-0.5004700	0.6062456
time__night	-0.1251296	0.8823825
time__nice	-0.0189295	0.9812485
lot__feel	-0.4591404	0.6318266
proud__happy	2.0159511	7.5078648
time__proud	1.4745939	4.3692612
food__dinner	-0.4055134	0.6666344
life__god	-0.5886247	0.5550902
time__im	0.8301088	2.2935682