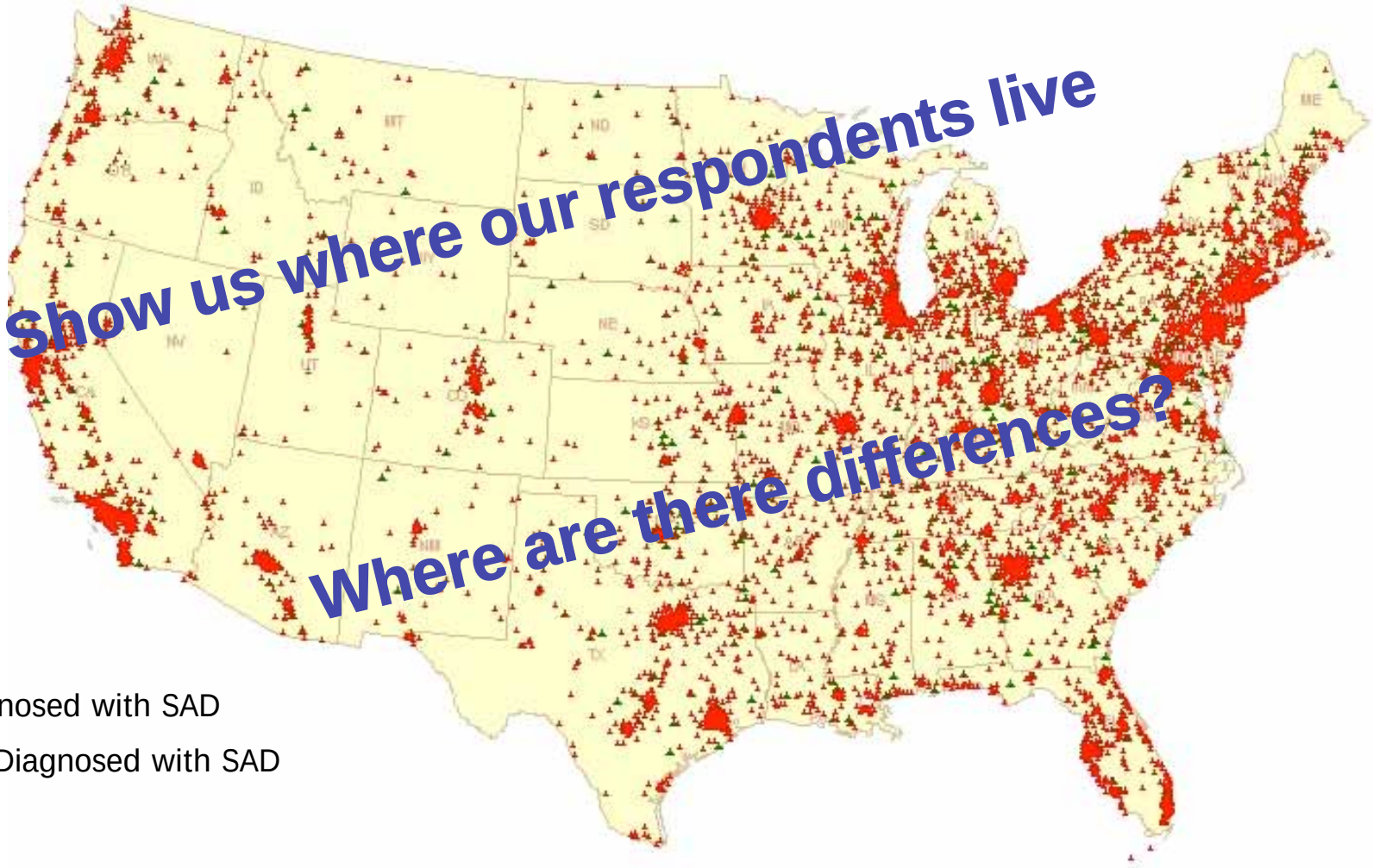
Using GIS to Confirm Latitudinal Evidence of Seasonal Affective Disorder (SAD)

July 27, 2005

What does Marketing Research have to do with GIS?

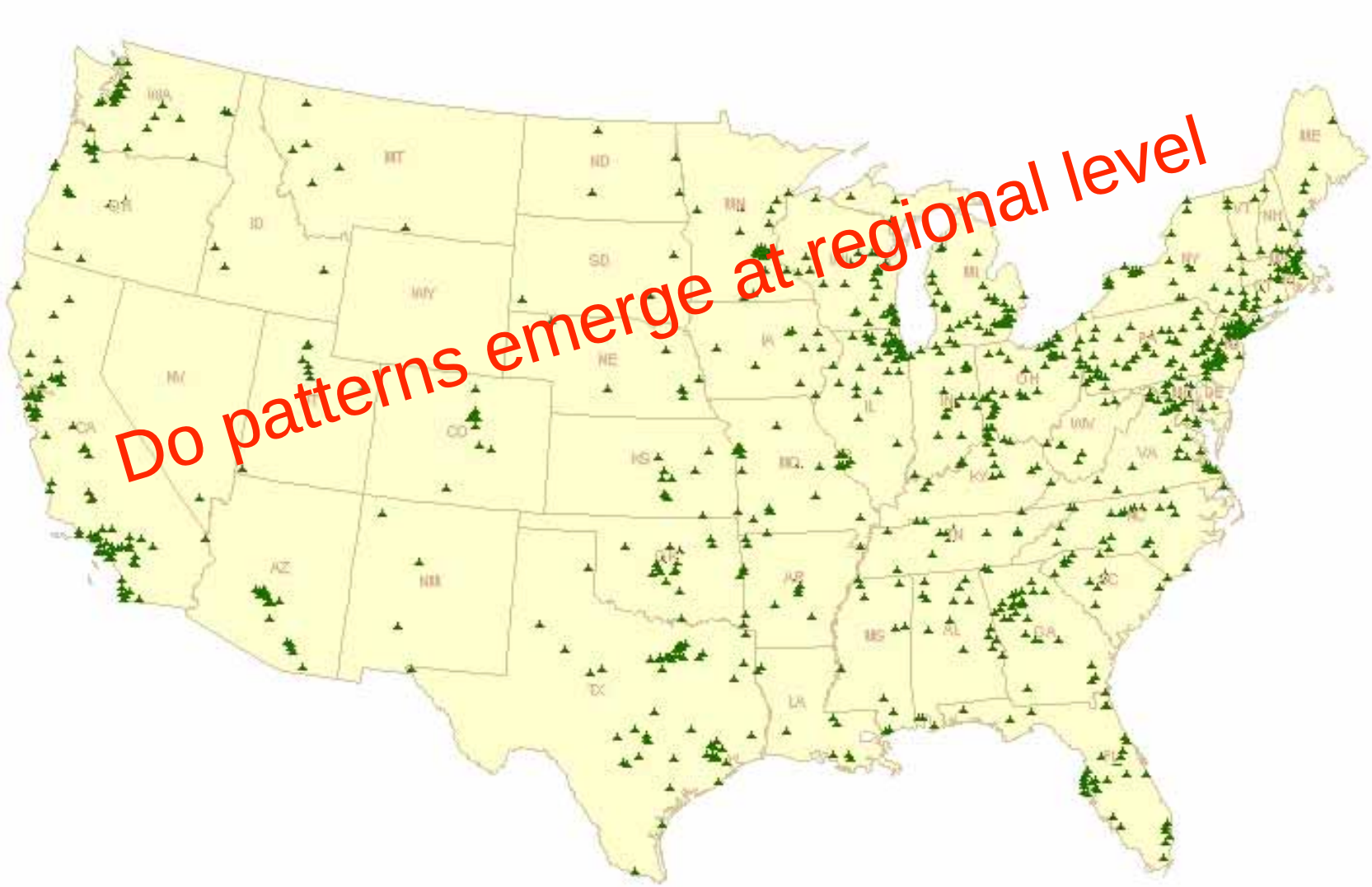
We use GIS to…….


Show us where our respondents live
Where are there differences?
Diagnosed with SAD
Not Diagnosed with SAD


Where are our groups of interest?

How do we know the map is telling a story worth listening to?

By making sure we base our story on a robust sample



Introduction

- Seasonal Affective Disorder (SAD) is a condition characterized by regularly recurring episodes of depression, usually in autumn and winter, alternating with periods of diminished or no depression during the spring and summer. SAD is also known as Major Depressive Disorder (MDD) with a seasonal pattern or, colloquially, as winter blues.

- Previous studies have indicated there is a positive relationship between latitude and the prevalence of SAD, with the highest prevalence at higher latitudes.

- The objective of this analysis was to assess the predictive value of selected geographic and climatologic variables on the prevalence of SAD in the 48 contiguous United States.

This study was funded by GlaxoSmithkline

Methodology

- Who:
  - 30,000 adults (over 18 years of age) were selected from the Ipsos-Insight NA panel
- How:
  - SAD was identified using standard diagnostic scoring
  - Each respondents zip code was linked to latitude and to regional climatologic data accessed through the National Climatic Data Center (Asheville, NC)
  - Regression analysis was used to assess the relationship between and unweighted prevalence of SAD and independent geographic and climatologic variables

An initial hypothesis was that SAD prevalence in the US would vary by geography

- To this end, we verified zip code information to permit analysis by region of the country, by state, and by major metropolitan area.

- We considered, but rejected, Census divisions as a meaningful way to analyze the prevalence data.
  - The grouping of states within Census division frequently obscured the prevalence data. This is because some divisions include states that span both northern and southern regions of the US.

- Our hypothesis was well-founded when looking at the data by ***LATITUDES***.

It was hypothesized that as the latitude increases, the incidence of SAD increases. A preliminary review of the data suggested that although the correlation between latitude and SAD was significant it was only negative .036.

Descriptives

			Statistic	Std. Error
SADo_1	Mean		.0742	.00228
	95% Confidence Interval for Mean	Lower Bound	.0698	
		Upper Bound	.0787	
	5% Trimmed Mean		.0269	
	Median		.0000	
	Variance		.069	
	Std. Deviation		.26218	
	Minimum		.00	
	Maximum		1.00	
	Range		1.00	
	Interquartile Range		.00	
	Skewness		3.248	.021
	Kurtosis		8.553	.043

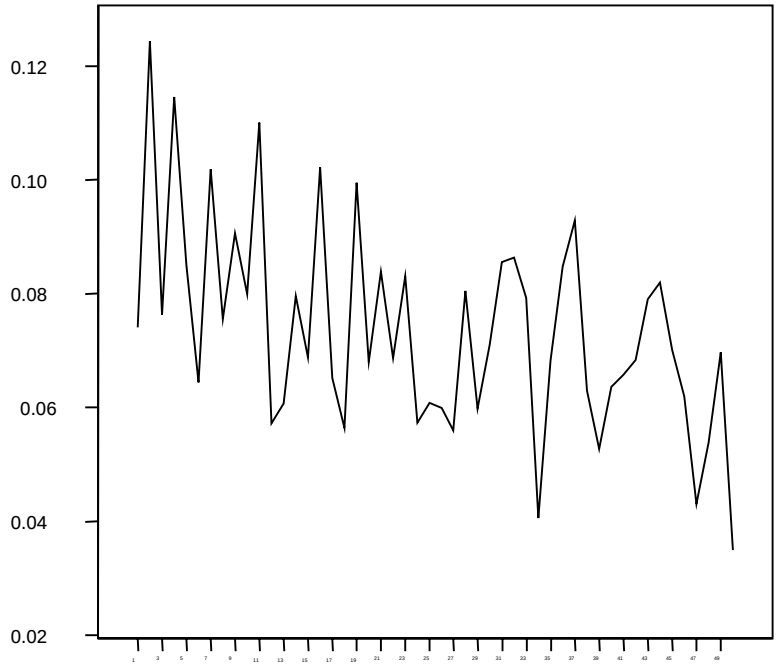
Correlations

		Latitude	sad
Latitude	Pearson Correlation	1	-.036**
	Sig. (2-tailed)	.	.000
	N	13173	13173
sad	Pearson Correlation	-.036**	1
	Sig. (2-tailed)	.000	.
	N	13173	13173

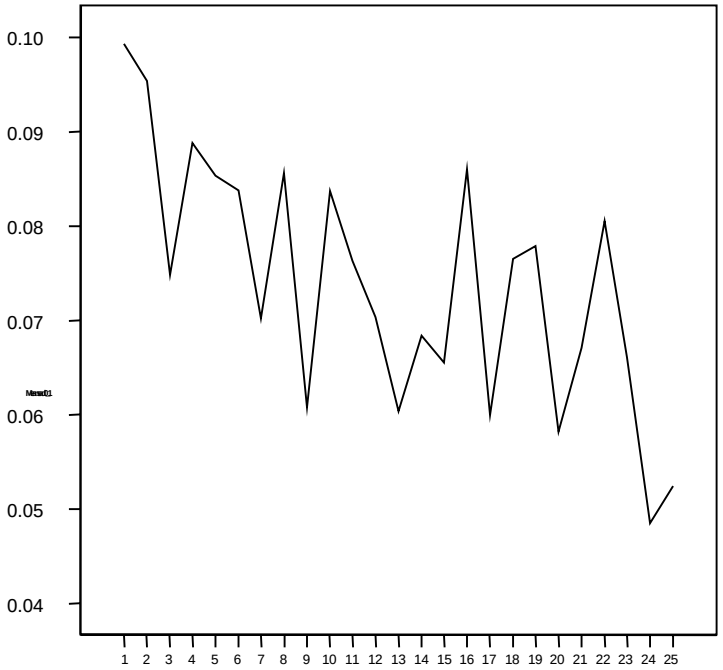
**. Correlation is significant at the 0.01 level

Note that the correlation between the latitude and the incidence of SAD is -.036, but significant.

The data were progressively grouped into larger and larger latitudinal regions, each region having similar numbers of respondents with SAD (by SPAQ).



50 NTILES of Latitude

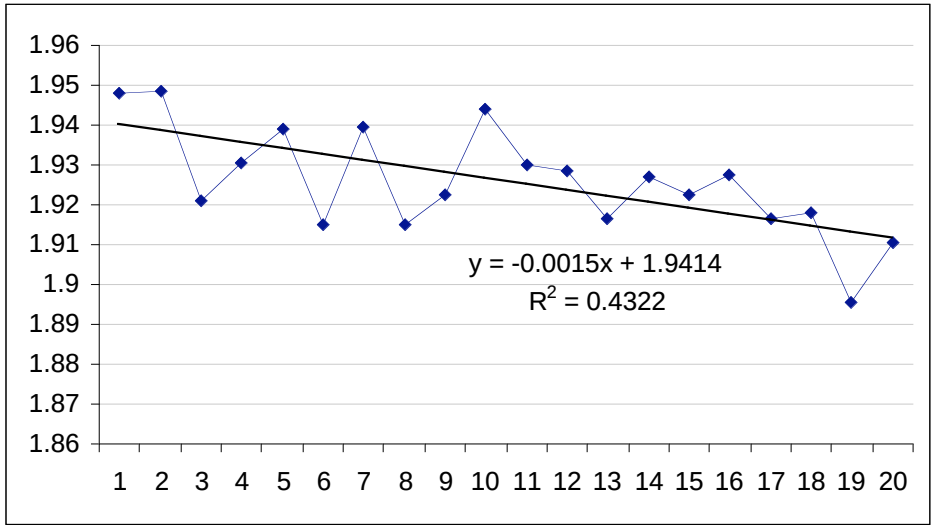


25 NTILES of Latitude

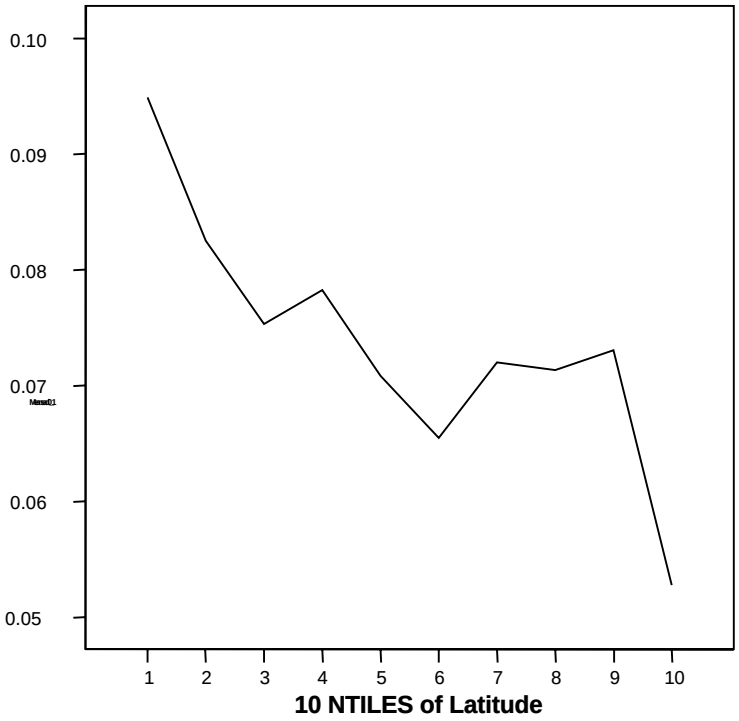
However, when latitudes were combined into 20 regions, the region by SAD incidence correlation became much higher, -.6574, suggesting that the hypothesis might have merit.

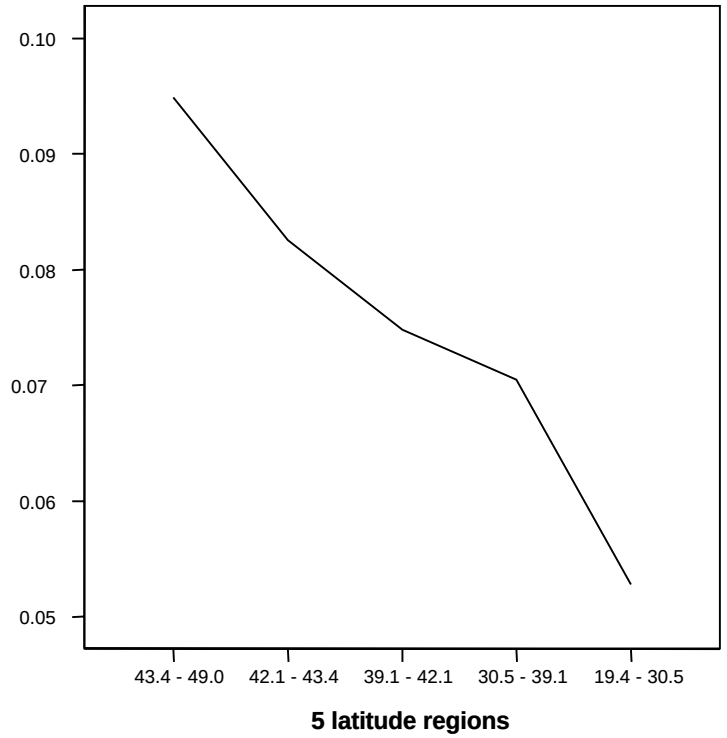
	Latitude			
REGION	Minimum	Maximum	Valid N	mean
1	19.44	29.15	N=657	1.948
2	29.16	30.51	N=659	1.948
3	30.51	32.84	N=659	1.921
4	32.84	33.63	N=661	1.930
5	33.63	34.06	N=657	1.939
6	34.06	35.11	N=659	1.915
7	35.11	36.11	N=659	1.939
8	36.11	37.54	N=659	1.915
9	37.54	38.57	N=657	1.922
10	38.57	39.10	N=660	1.944
11	39.10	39.76	N=658	1.930
12	39.76	40.29	N=658	1.929
13	40.30	40.74	N=660	1.917
14	40.74	41.08	N=659	1.927
15	41.09	41.64	N=657	1.922
16	41.64	42.10	N=660	1.927
17	42.11	42.62	N=658	1.916
18	42.63	43.37	N=659	1.918
19	43.38	45.11	N=659	1.895
20	45.12	48.96	N=658	1.910

Note that the correlation between the region and the incidence of SAD is now -.6574.



When respondents were pooled into five latitudinal regions, the approximately linear relationship between SAD prevalence and latitude was strongest





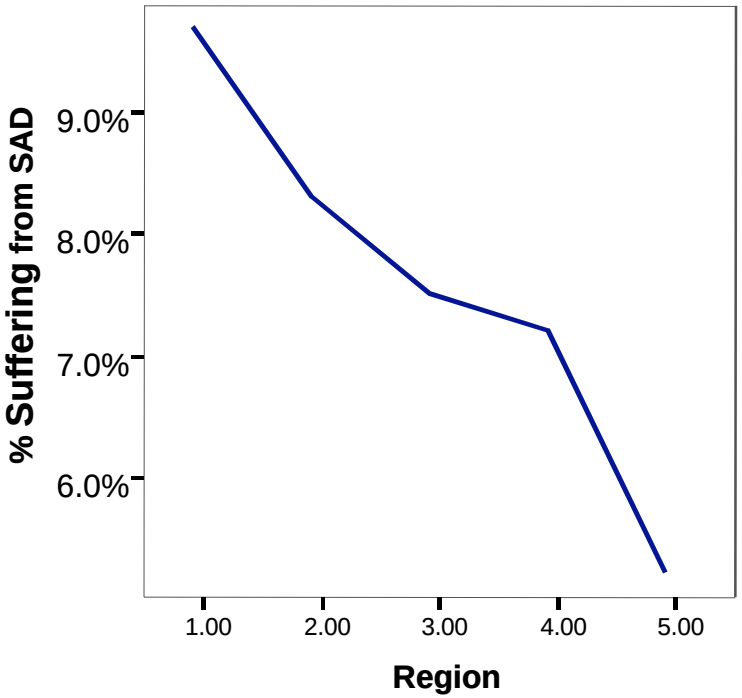
Regional data from the table at the left was combined into five latitude regions, summarized in the table on the right, with the objective of having at least 30 SAD suffers in any given region.

region	Latitude Range			SAD sufferers	
20	45.12	48.96	N=658	1.91	9.0%
19	43.38	45.11	N=659	1.9	10.5%
18	42.63	43.37	N=659	1.92	8.2%
17	42.11	42.62	N=658	1.92	8.4%
16	41.64	42.10	N=660	1.93	7.3%
15	41.09	41.64	N=657	1.92	7.8%
14	40.74	41.08	N=659	1.93	7.3%
13	40.30	40.74	N=660	1.92	8.3%
12	39.76	40.29	N=658	1.93	7.1%
11	39.10	39.76	N=658	1.93	7.0%
10	38.57	39.10	N=660	1.94	5.6%
9	37.54	38.57	N=657	1.92	7.8%
8	36.11	37.54	N=659	1.92	8.5%
7	35.11	36.11	N=659	1.94	6.1%
6	34.06	35.11	N=659	1.92	8.5%
5	33.63	34.06	N=657	1.94	6.1%
4	32.84	33.63	N=661	1.93	7.0%
3	30.51	32.84	N=659	1.92	7.9%
2	29.16	30.51	N=659	1.95	5.2%
1	19.44	29.15	N=657	1.95	5.2%

	LATITUDE		suffering from SAD			not suffering from SAD			Total
region	Minimum	Maximum	Count	Row %	Col %	Count	Row %	Col %	Count
1	43.38	48.96	128	9.7	13.1	1189	90.3	9.7	1317
2	42.11	43.37	109	8.3	11.1	1208	91.7	9.9	1317
3	39.10	42.10	295	7.5	30.2	3657	92.5	30.0	3952
4	30.51	39.10	378	7.2	38.7	4893	92.8	40.1	5271
5	19.44	30.51	68	5.2	7.0	1248	94.8	10.2	1316
		Total	978	7.4	100.0	12195	92.6	100.0	13173

Prevalence of SAD and Latitudes

Respondent data grouped into five latitudinal regions shows a linear relationship between SAD prevalence and latitude: y = -1.029x + 10.622 (R2 = 0.9423)



LATITUDE			SAD Dx			Not SAD Dx			Total
Reg	Min	Max	Count	Row %	Col %	Count	Row %	Col %	Count
1	43.3764	48.9552	128	9.7	13.1	1189	90.3	9.7	1317
2	42.1072	43.3671	109	8.3	11.1	1208	91.7	9.9	1317
3	39.0985	42.1036	295	7.5	30.2	3657	92.5	30.0	3952
4	30.5137	39.0959	378	7.2	38.7	4893	92.8	40.1	5271
5	19.4393	30.5126	68	5.2	7.0	1248	94.8	10.2	1316
		Total	978	7.4	100.0	12195	92.6	100.0	13173

Summary

Latitude Analysis

- It was hypothesized that as the latitude increases, the prevalence of SAD increases
- A preliminary review of the data suggested that although the correlation between latitude and SAD was significant it was only -.036
- However, when latitudes were combined into 20 regions, the region by SAD prevalence correlation became much higher, -.6574, suggesting that the hypothesis might have merit
- Regional data from the 20 region analysis was combined into 5 latitude regions, with the objective of having at least 30 SAD sufferers in any given region
- As the latitude increases from 19 in the South of the USA to 49 in the North, the prevalence of SAD in the sample (13,173) increased from 5.2% to 9.7%. Of interest is that this relationship is linear across the five regions suggesting that the prevalence declines by about 1% per region from North to South:

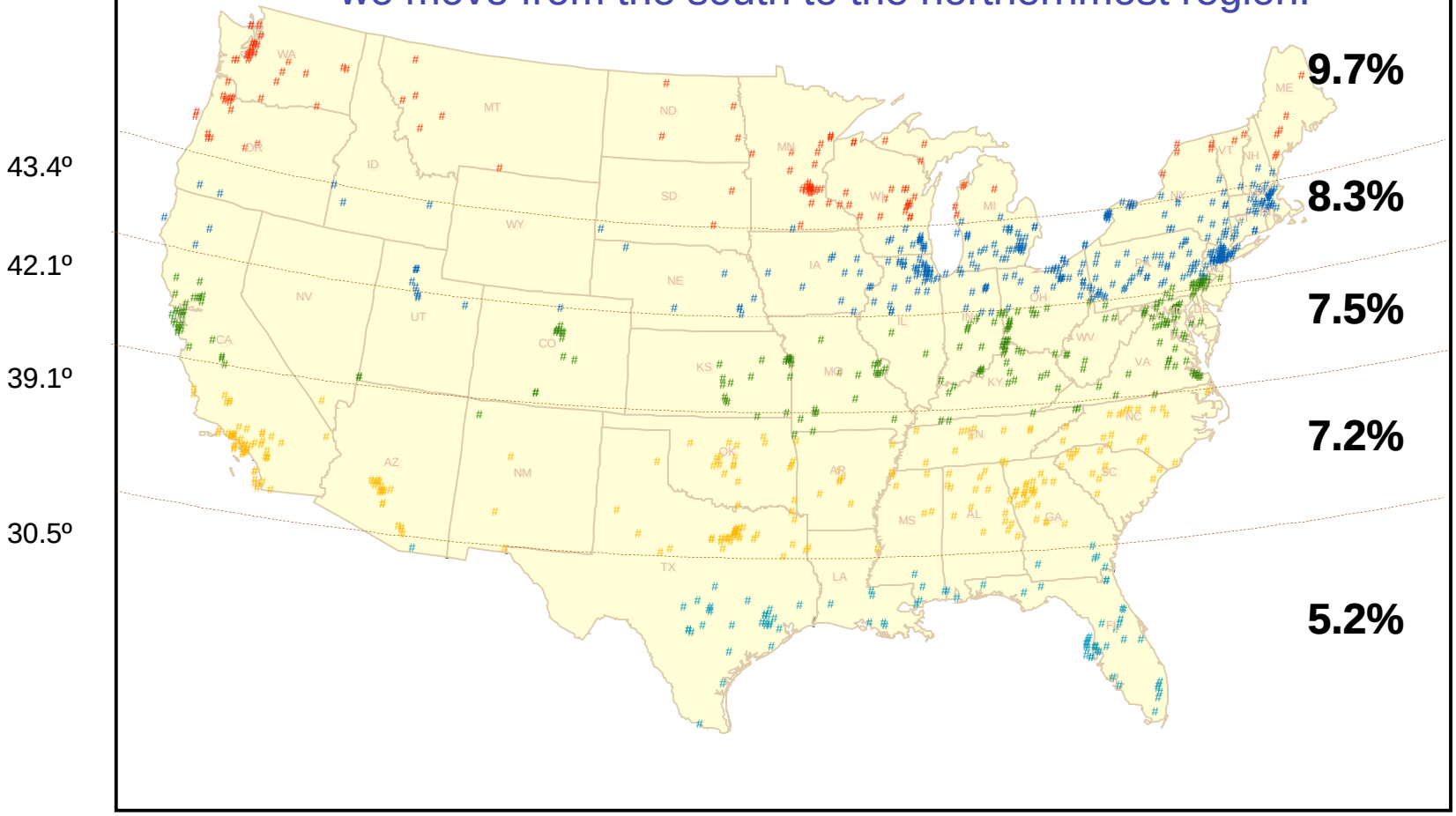
y = -1.029x + 10.622 (R-sq = .9423)

Robust sample sizes within each latitude region to make inferences that are statistically significant

Legend	% Respondents	Latitudes
▲	9.7%	43.4-48.9
▲	8.3%	42.1-43.4
▲	7.5%	39.1-42.1
▲	7.2%	30.5.-39.1
▲	5.2%	19.4-30.5

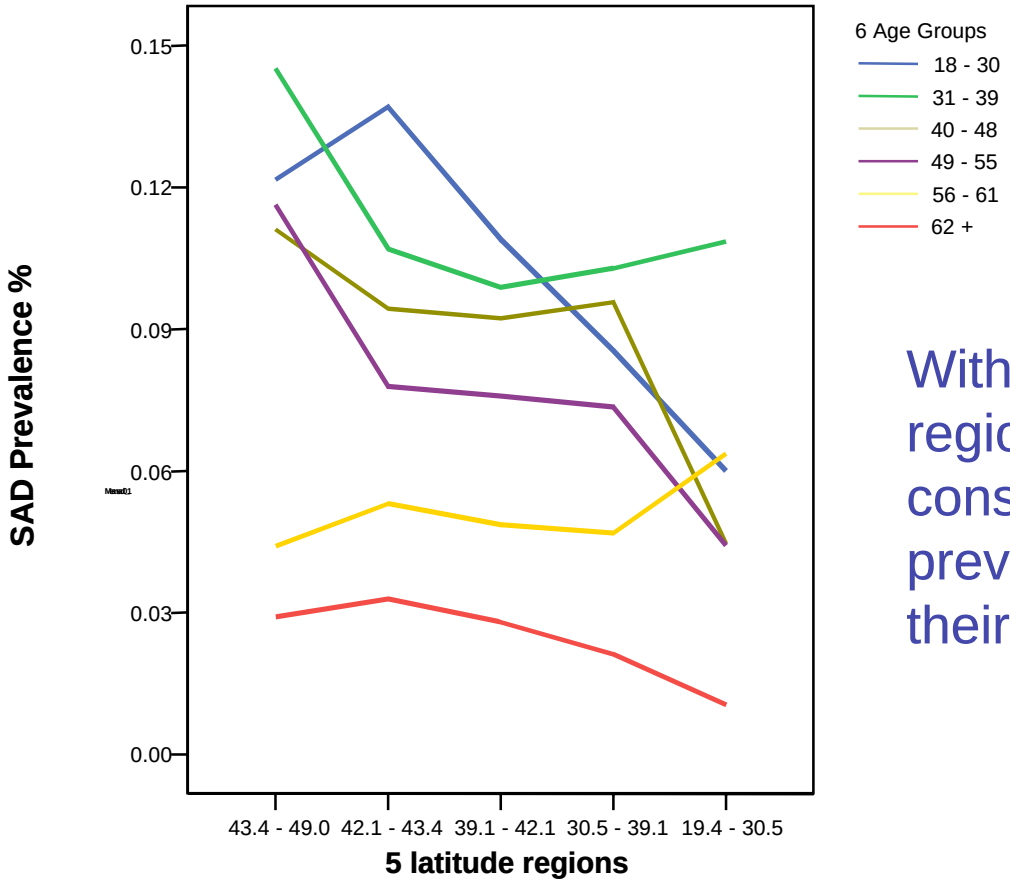
Latitudes on the map present a clear picture

- The prevalence of SAD increases by 1% by latitude region, as we move from the south to the northernmost region.



Having respondent level data helped us look at our findings by different demographic groups of interest

Prevalence of SAD by Age Groups within Latitude Regions



Within each latitude region, older people consistently have lower prevalence of SAD than their younger cohorts.

North **South**

Seasonal variation in daylight hours is also highly predictive of SAD prevalence

- SAD prevalence increases as the seasonal variance in hours of daylight increases.
- The greater the change in hours of daylight, from summer to winter, the higher the prevalence of SAD.

Using the average number of hours of daylight for the latitudinal regions is not interesting because the seasonal variance averages out to roughly 12:00 hours of daylight. What is of interest is the daylight variance by latitude.

At the equator, there is no variance, but in the latitudes of interest, the seasonal daylight variance ranges from 1:10 minutes in the South to nearly 9 hours in the North, in other words at 50 degrees North, daylight on May 15th is 17:04 and shortens to 8:05 on December 15th, a change of 9 hours.

Length of Day at Various Latitudes (in Hours and Minutes on the 15th of Each Month)

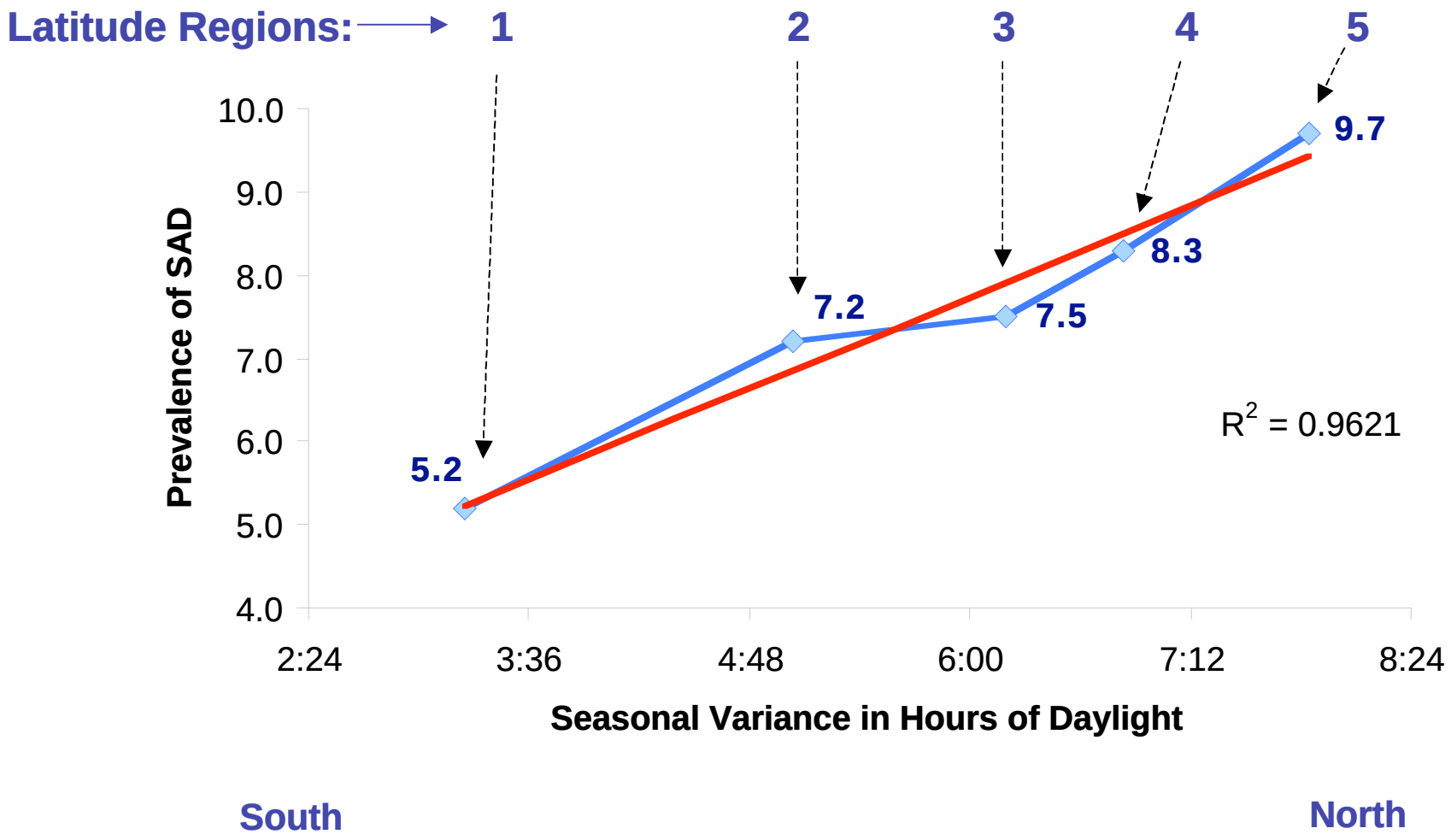
Month	*Equator*	*10°*	*20°*	*30°*	*40°*	*50°*	*60°*	*70°*	*80°*	*Poles*	*Month*
January	12:07	11:35	11:02	10:24	9:37	8:30	6:38	0:00	0:00	0:00	July
February	12:07	11:49	11:21	11:10	10:42	10:07	9:11	7:20	0:00	0:00	August
March	12:07	12:04	12:00	11:57	11:53	11:48	11:41	11:28	10:52	0:00	September
April	12:07	12:21	12:36	12:53	13:14	13:44	14:31	16:06	24:00:00	24:00:00	October
May	12:07	12:34	13:04	14:22	15:22	17:04	22:13	24:00:00	24:00:00	24:00:00	November
June	12:07	12:42	13:20	14:04	15:00	16:21	18:49	24:00:00	24:00:00	24:00:00	December
July	12:07	12:40	13:16	13:56	14:49	15:38	17:31	24:00:00	24:00:00	24:00:00	January
August	12:07	12:28	12:50	13:16	13:48	14:33	15:46	18:26	24:00:00	24:00:00	February
September	12:07	12:12	12:17	12:23	12:31	12:42	13:00	13:34	15:16	24:00:00	March
October	12:07	11:55	11:42	11:28	11:10	10:47	10:11	9:03	5:10	0:00	April
November	12:07	11:40	11:12	10:40	10:01	9:06	7:37	3:06	0:00	0:00	May
December	12:07	11:32	10:56	10:14	9:20	8:05	5:54	0:00	0:00	0:00	June

Average Daylight	12:07	12:08	12:13	12:17	12:22

Daylight Variance	1:10	2:24	4:08	6:02	8:59

Prevalence of SAD by Seasonal Variance in Hours of Daylight

Within 5 Latitude regions



Cloudy Day Analysis

- A grouping of the cloudy states to increase the sample size results in a significant correlation of .328

Correlations

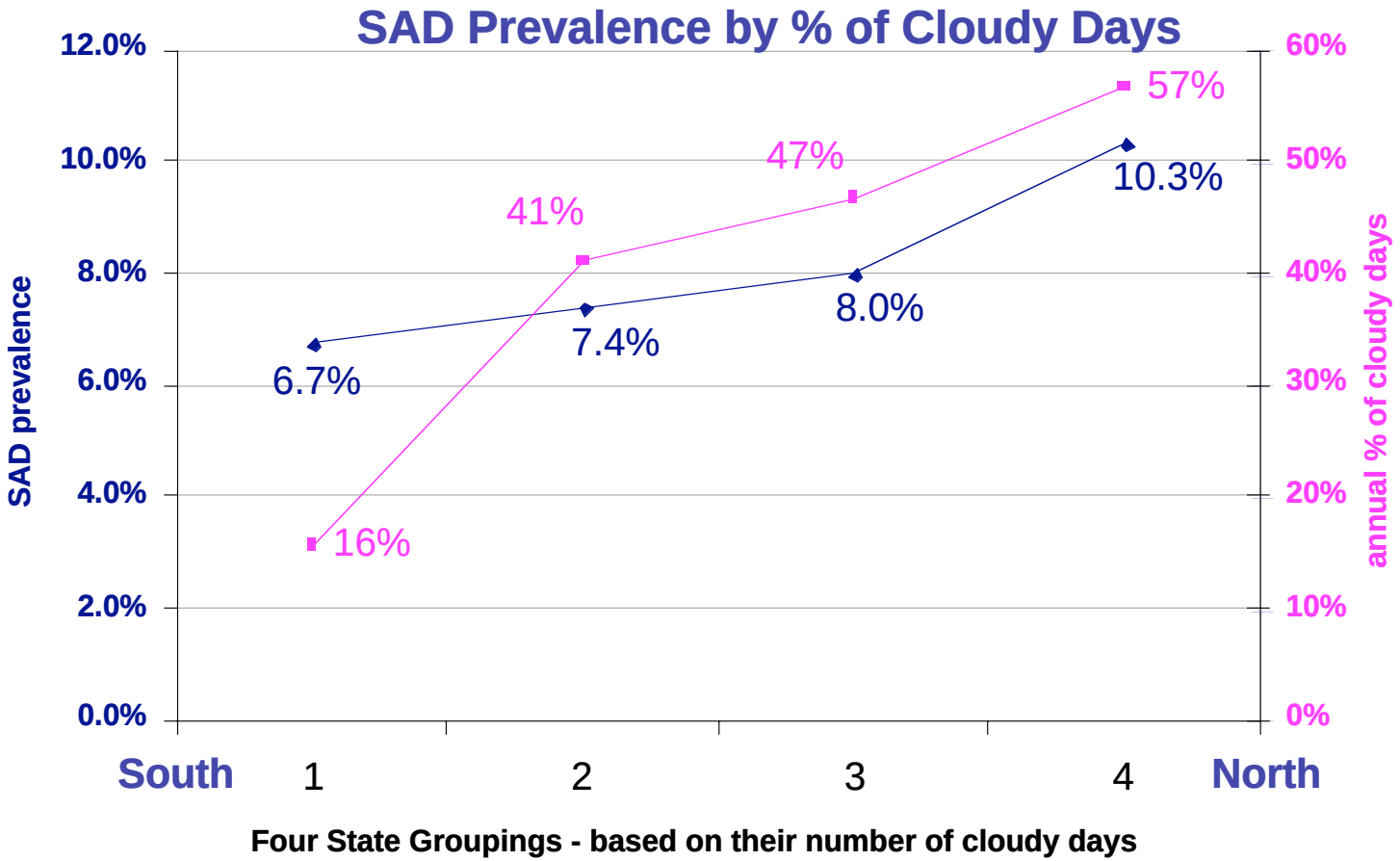
		cloudy	SAD	NTILES of cloudy
cloudy	Pearson Correlation	1	.277	.757**
	Sig. (2-tailed)	.	.054	.000
	N	49	49	49
SAD	Pearson Correlation	.277	1	.328*
	Sig. (2-tailed)	.054	.	.021
	N	49	49	49
NTILES of cloudy	Pearson Correlation	.757**	.328*	1
	Sig. (2-tailed)	.000	.021	.
	N	49	49	49

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

The more the cloudy days, the greater the prevalence of SAD

- Cloudy days are predictive of SAD. It is important to note that most of the states with a lot of cloudy days are in the northern US.



State Groups

Used in Cloudy Day Analysis

State	SAD prevalence	cloudy days	group
AZ	6%	0.22	1
CA	6%	0.29	1
CO	6%	0.24	1
FL	5%	0.32	1
KS	7%	0.26	1
NM	4%	0.25	1
OK	12%	0.29	1
TX	6%	0.32	1

State	SAD prevalence	cloudy days	group
AL	10%	0.42	2
AR	11%	0.4	2
CT	5%	0.46	2
DC	3%	0.45	2
DE	8%	0.45	2
GA	9%	0.41	2
IL	7%	0.46	2
LA	7%	0.38	2
MA	10%	0.45	2
MD	7%	0.42	2
MO	9%	0.44	2
MS	10%	0.36	2
NC	6%	0.42	2
ND	14%	0.45	2
NE	7%	0.35	2
NJ	5%	0.44	2
NV	2%	0.32	2
NY	7%	0.45	2
RI	2%	0.44	2
SC	8%	0.41	2
SD	11%	0.42	2
TN	5%	0.43	2
UT	12%	0.34	2
VA	7%	0.42	2
WY	0%	0.39	2

State	SAD prevalence	cloudy days	group
IA	7%	0.46	3
ID	6%	0.47	3
IN	7%	0.49	3
KY	8%	0.48	3
MN	11%	0.49	3
OR	6%	0.5	3
PA	9%	0.49	3
WI	11%	0.48	3

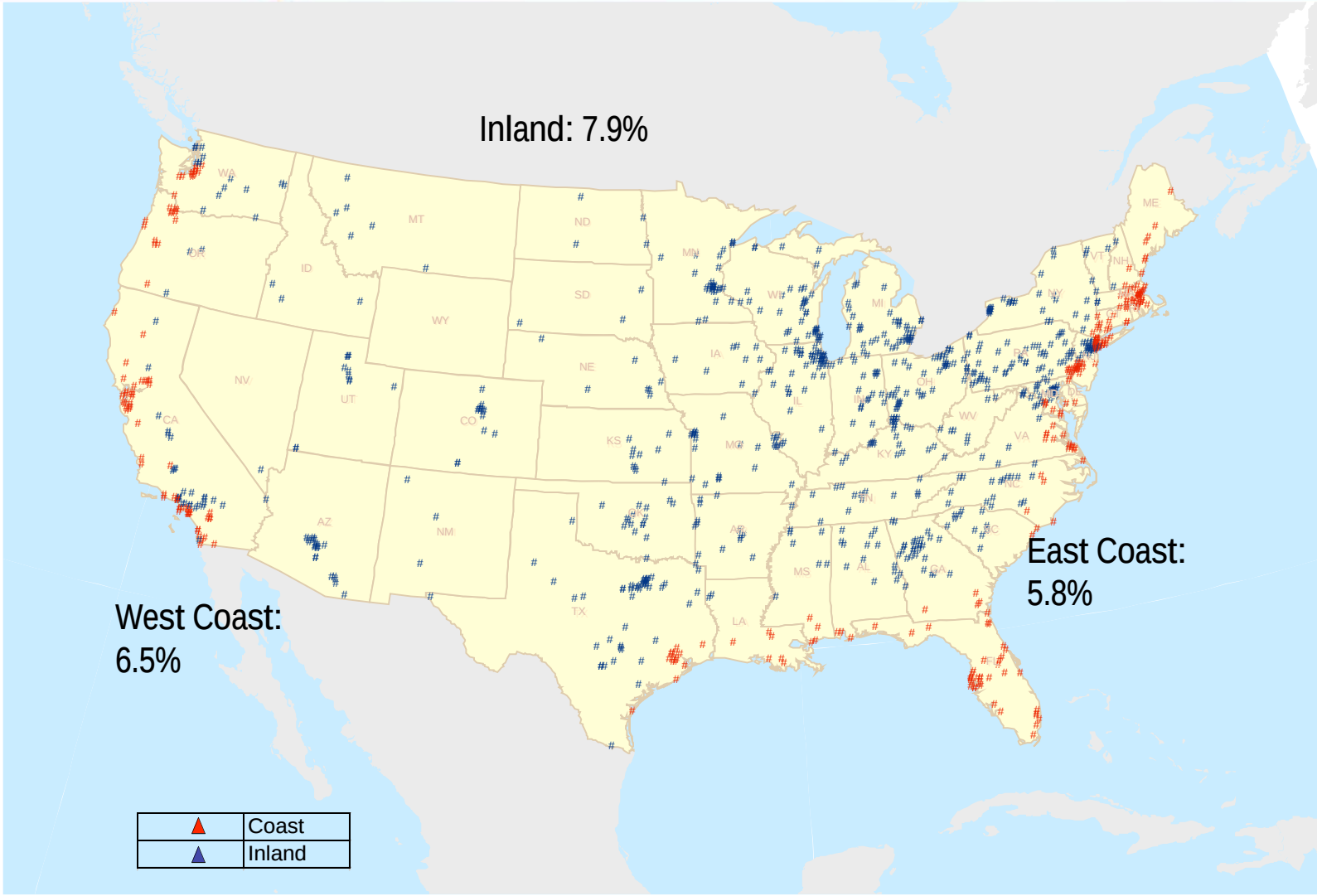
State	SAD prevalence	cloudy days	group
ME	9%	0.51	4
MI	10%	0.54	4
MT	14%	0.51	4
NH	9%	0.56	4
OH	9%	0.53	4
VT	16%	0.56	4
WA	10%	0.57	4
WV	9%	0.56	4

Interestingly, people living in the coastal regions have a lower prevalence of SAD that those living further inland

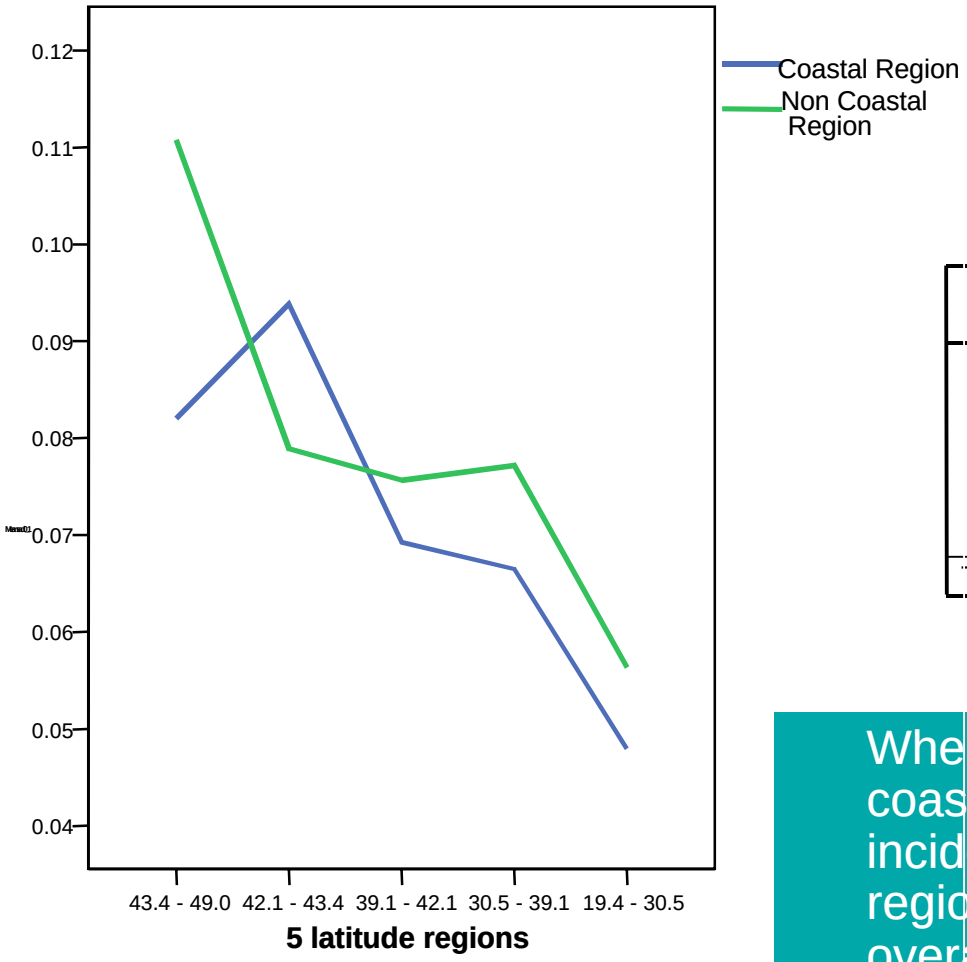
- However, within a single coastal region (e.g., the Pacific Coast census division), SAD prevalence varies dramatically by latitude from South to North.

- On the Pacific Coast, prevalence ranges…
  - From a low of 1.6% in the southernmost latitude region (southern California)…
  - To a high of 8% in the most northern latitude of the Pacific Coast (Washington State)
  - Controlling for latitude, would there be a coastal effect?

Further research is required to determine if this significant difference in SAD prevalence between coastal & inland is confounded by other factors like the size of the city where the respondent lives.



SAD vs. Non-SAD by Coastal Regions



SAD vs. Non-SAD by Coastal Regions

		sad		
		1.00	2.00	Total
Coastal Region	Count	342	4771	5113
	Adjusted Residual	-2.4	2.4	
Non Coastal Region	Count	643	7602	8245
	Adjusted Residual	2.4	-2.4	
Total	Count	985	12373	13358

When latitude is taken into consideration, coastal regions do appear to have a lower incidence of SAD than do non coastal regions. The cross tabulation indicates an overall significant relationship at the 99%

Conclusions

- This is the first nationwide study to confirm the direct relationship between the prevalence of SAD and geographic latitude among adults living in the United States. The results of this study were presented at the American Psychiatric Association conference in May 2005.

- The prevalence of SAD was highest in the northern latitudes and it declined with decreasing latitudes. However, SAD remains relatively common even in the southern latitudes of the United States.

- Within this study, the strength of the relationship between the prevalence of SAD and latitude appears to be derived principally from the seasonal variation in the hours of daylight. From this study, the magnitude of this seasonal change was most strongly correlated with the prevalence of SAD.

Conclusions

- The results of this study suggest that the quality of daylight is important. Annual average percent of cloudy days was directly correlated with the prevalence of SAD. However, annual percent of cloudy days is not independent of latitude and there is a higher average annual percent of cloudy days in the northern latitudes compared to the southern latitudes.

- Even when latitude is taken into consideration, it appears the prevalence of SAD is slightly lower among those who live in coastal areas when compared to those who live in non-coastal areas. This relationship was evident, even though the average annual percent of cloudy days is higher in the coastal areas when compared to non-coastal areas. This relationship may deserve further investigation.

- Regional differences are best illustrated with the use of GIS maps