PENNSTATE

The Future of Intro Stat: More Accessible, More Effective, More Relevant, More Fun!

Kari Lock Morgan
Department of Statistics
Penn State University

STATS 2020
Baltimore, MD
November 22nd, 2019

There is a revolution going on in this course, across the country and across all types of institutions

Driven by

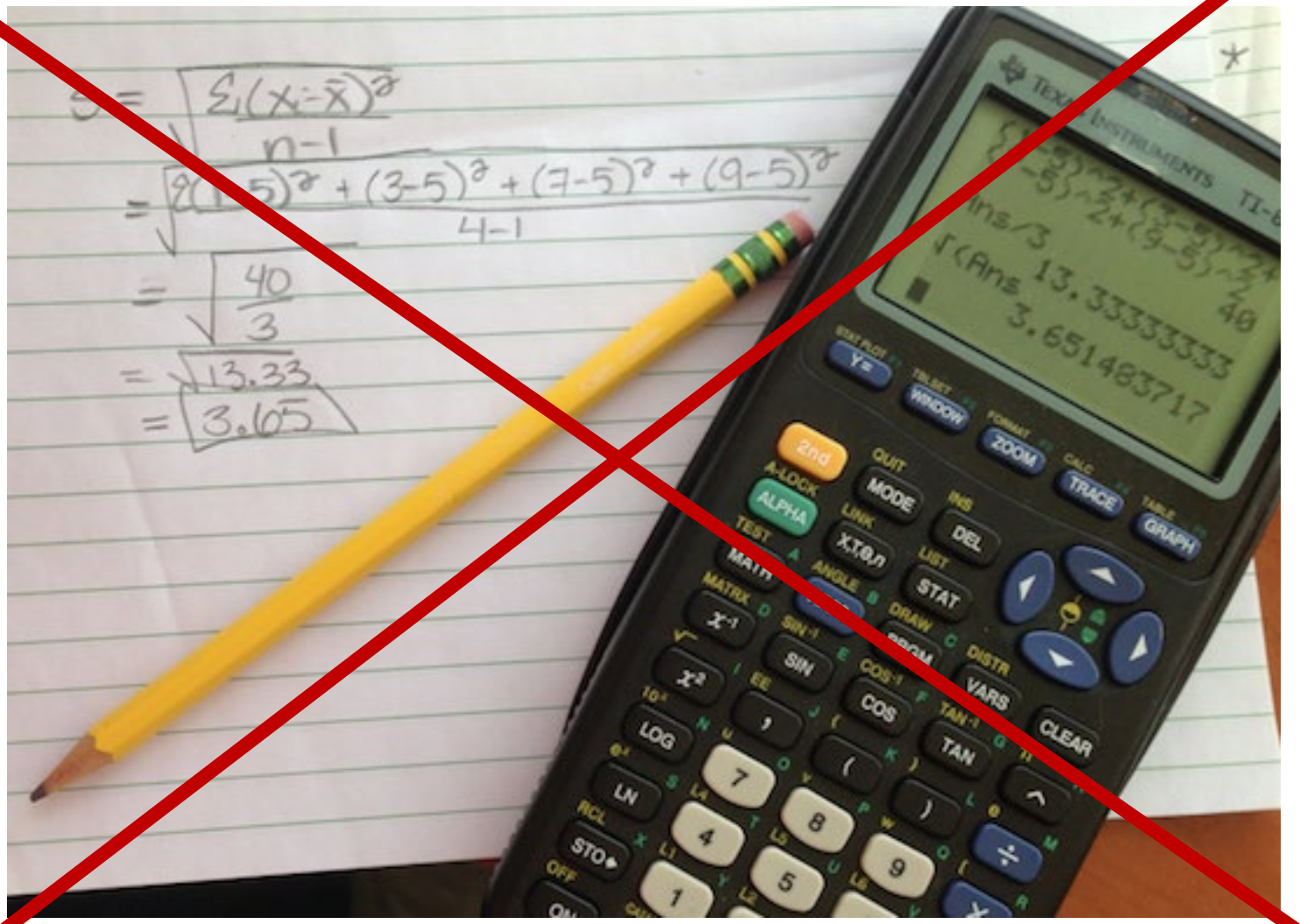
TECHNOLOGY!!!

Potential to dramatically enhance student enjoyment and student success!

Intro Stats: Not This

Consider two events A and B, and assume that P(A) = 0.6 and P(B) = 0.5 and $P(A \cap B) = 0.2$.
Find $P(A \cup B)$.

Intro Stats: Not This



Intro Stats: Not This

	Confidence Interval	Test Statistic
General	Sample statistic $\pm z^* \cdot SE$	$\frac{\text{Sample statistic} - \text{Null parameter}}{SE}$
Proportion	$\hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$\frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
Mean	$\bar{x} \pm t^* \cdot s/\sqrt{n}$	$\frac{\bar{x} - \mu_0}{s/\sqrt{n}}$
Difference in Proportions	$(\hat{p}_1 - \hat{p}_2) \pm z^* \cdot \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$	$\frac{(\hat{p}_1 - \hat{p}_2) - 0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}}$
Difference in Means	$(\bar{x}_1 - \bar{x}_2) \pm t^* \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
Paired Difference in Means	$\bar{x}_d \pm t^* \cdot \frac{s_d}{\sqrt{n_d}}$	$\frac{\bar{x}_d - 0}{s_d/\sqrt{n_d}}$

Intro Stats: But This

Are mosquitoes more attracted to beer drinkers?
Are the youngest kids in a class more likely to be diagnosed with ADHD?
Does leaving a light on at night affect weight?
Is there a "commitment" gene?
Does diet cola leach calcium out of the system?
Does drinking red wine boost metabolism?
Do[illegible]s ra[illegible]ia[illegible]on fro[illegible]ll p[illegible]ones aff[illegible]ct b[illegible]in a[illegible]ity?
Ar[illegible]e[illegible]s m[illegible]re lik[illegible]t[illegible]att[illegible]ck after[illegible] full[illegible]n[illegible]?
Do people read faster from a printed book than from an iPad or Kindle?
Does tagging penguins for identification purposes harm them?
Does turning up the music in a bar increase beer consumption?
Are city dwellers more likely to have mood and anxiety disorders?
Can dogs smell cancer in humans?
Does sexual frustration increase the desire for alcohol?
How broadly do experiences of parents affect their future children?
What percent of college professors consider themselves "above average" teachers?
AND SO ON!

REAL DATA!!!

What's Out?

~~Computing statistics by hand~~

~~Drawing graphs by hand~~

~~Probability rules~~

~~Discrete probability distributions~~

~~Abstract random variables~~

What's In?

2016 ASA GAISE Guidelines:

1. Teach **statistical thinking**.
2. Focus on **conceptual understanding**.
3. Integrate **real data** with a context and purpose.
4. Foster **active learning**.
5. Use **technology** to explore concepts and analyze data.
6. Use **assessments** to improve and evaluate student learning.

Ways to Change

1. Make everything about real **data**

Focus on Real Data

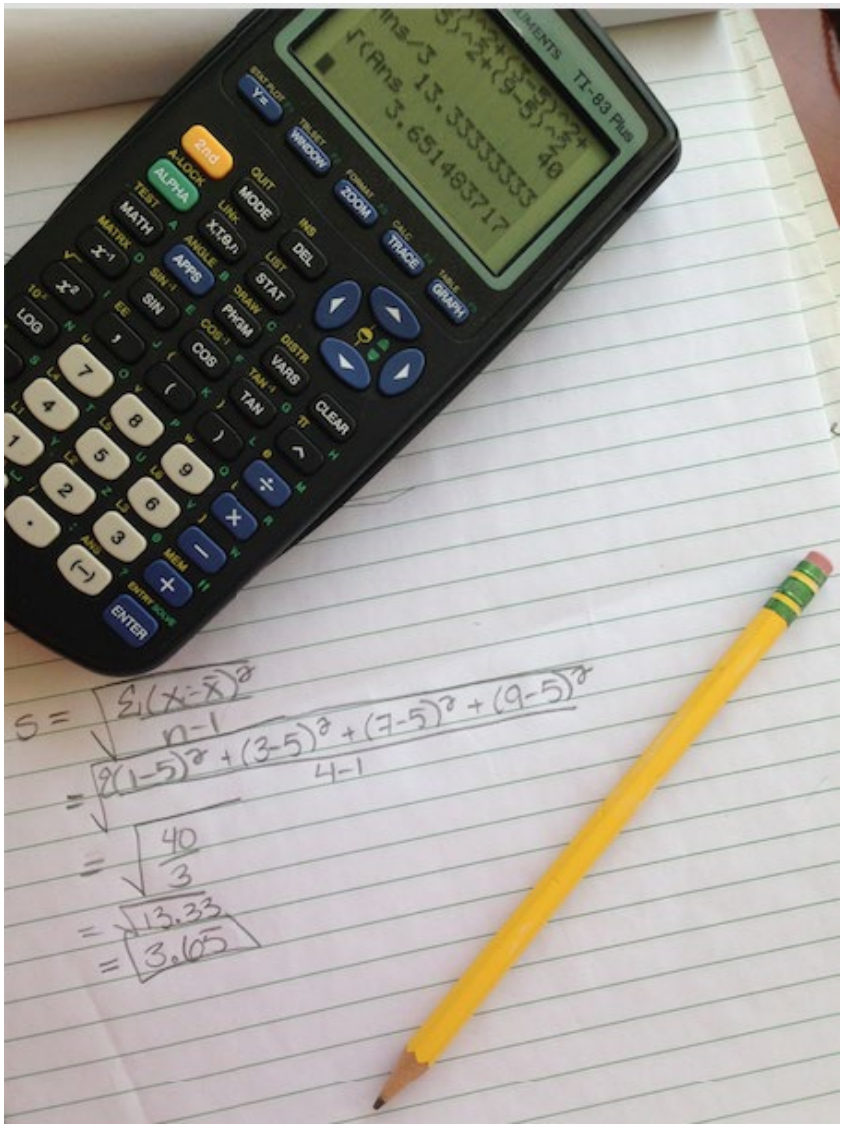
Topics of the day:

- Two-way tables
- Graphical displays for two categorical variables
- Difference in proportions

Ways to Change

1. Make everything about real **data**
2. Reduce tedious "by-hand" work and let **technology** do the heavy-lifting

Rely on technology



STATISTICAL SOFTWARE!!

Larger, real datasets!

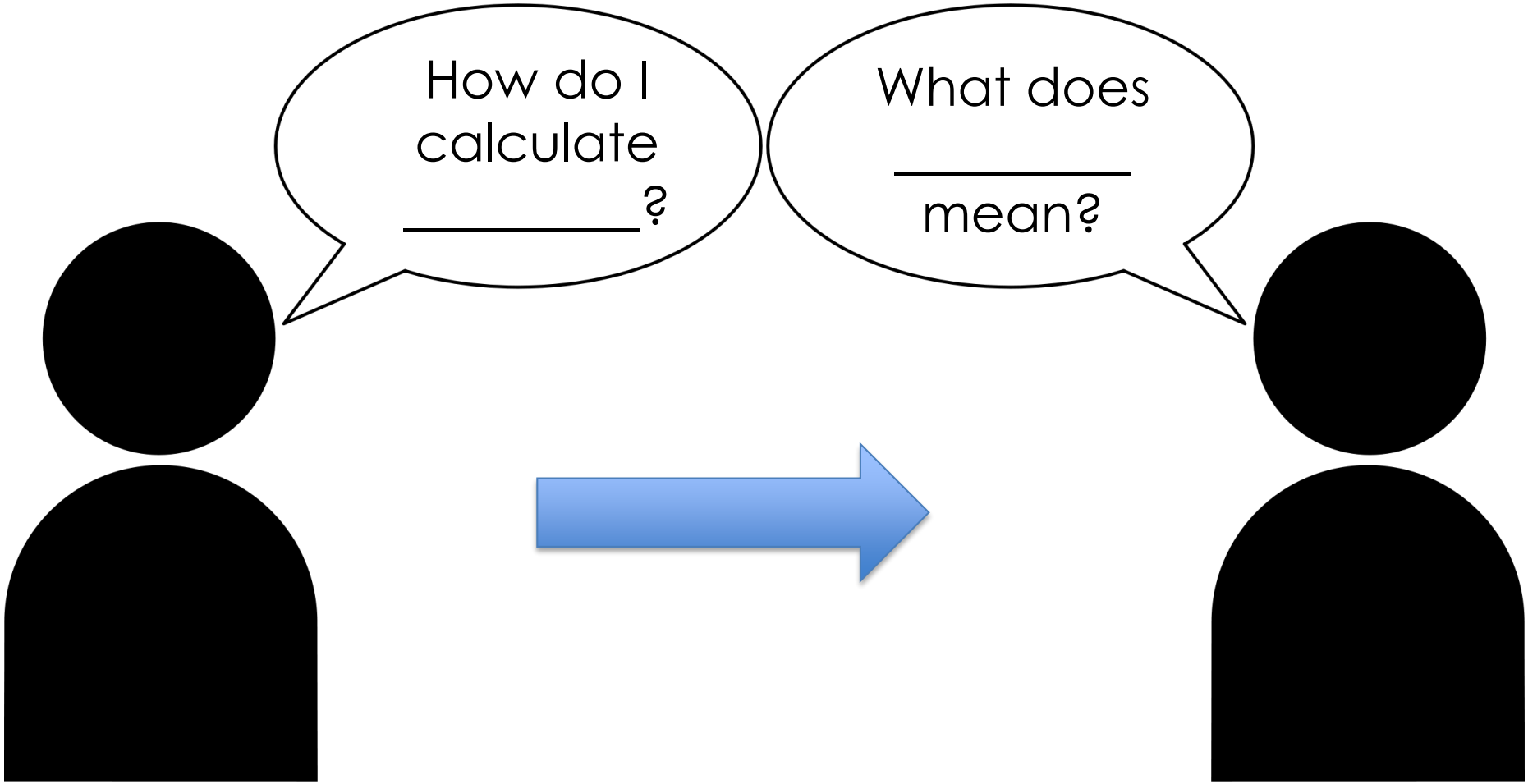
More exciting!

Interpretation!

Ways to Change

1. Make everything about real **data**
2. Reduce tedious "by-hand" work and let **technology** do the heavy-lifting
3. Focus on interpretation and **concepts**

Interpretation and concepts



Ways to Change

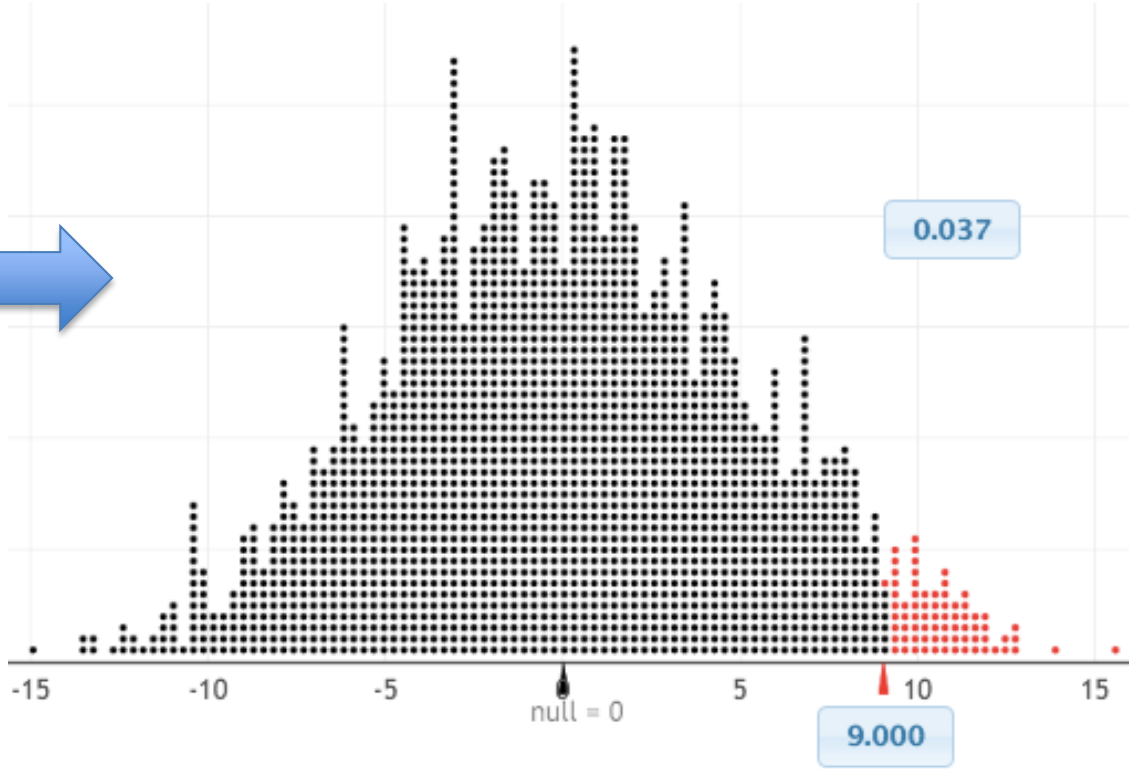
1. Make everything about real **data**
2. Reduce tedious "by-hand" work and let **technology** do the heavy-lifting
3. Focus on interpretation and **concepts**
4. Improve conceptual understanding and reduce reliance on prerequisites with **simulation-based inference**

Simulation-Based Inference

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

TABLE B: *t*-DISTRIBUTION CRITICAL VALUES

	Tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											



Example: Beer & Mosquitoes

Question: Does consuming beer attract mosquitoes?

Experiment:
25 volunteers drank a liter of beer,
18 volunteers drank a liter of water
Randomly assigned!
Mosquitoes were caught in traps as they approached the volunteers.[1]

[1] Lefvre, T., et. al., "Beer Consumption Increases Human Attractiveness to Malaria Mosquitoes," *PLoS ONE,* 2010; 5(3): e9546.

Beer and Mosquitoes

Number of Mosquitoes

Beer	Water
27	21
20	22
21	15
26	12
27	21
31	16
24	19
19	15
23	24
24	19
28	23
19	13
24	22
29	20
20	24
17	18
31	20
20	22
25	
28	
21	
27	
21	
18	
20	

Beer mean = 23.6

Water mean = 19.22

Beer mean – Water mean = 4.38

Two possible explanations:

- Beer attracts mosquitos
- No difference; random chance

Does drinking beer actually attract mosquitoes or is the difference just due to random chance?

Beer and Mosquitoes

Number of Mosquitoes

Beer	Water
27	21
20	22
21	15
26	12
27	21
31	16
24	19
19	15
23	24
24	19
28	23
19	13
24	22
29	20
20	24
17	18
31	20
20	22
25	
28	
21	
27	
21	
18	
20	

What kinds of results would we see, just by random chance, if there were no difference between beer and water?

We can simulate to find out!!!

Beer and Mosquitoes

Number of Mosquitoes

Beer					Water
27	27	19	21	24	21
20	20	24	18	19	22
21	21	29	20	23	15
26	26	20	21	13	12
27	27	27	22	22	21
31	31	31	15	20	16
24	24	20	12	24	19
19	19	25	21	18	15
23	23	28	16	20	24
24	24	21	19	22	19
28	28	27	15		23
19					13
24					22
29					20
20					24
17					18
31					20
20					22
25					
28					
21					
27					
21					
18					
20					

1. Assume no difference (beer/water doesn't matter).
2. Mimic random chance: Re-randomize the 43 values into two groups of 25 and 18

Beer and Mosquitoes

Number of Mosquitoes

Beer			Water
27	19	21	24
20	24	18	19
24	29	20	23
16	20	31	13
27	27	19	22
31	31	15	20
24	20	18	24
19	25	22	18
28	28	16	20
24	21	14	22
28	27	15	
21		20	
18		27	
15		21	
21		17	
16		24	
28		28	
22			
19			
27			
20			
23			
22			
21			

1. Assume no difference (beer/water doesn't matter).
2. Mimic random chance: Re-randomize the 43 values into two groups of 25 and 18
3. Compute the beer mean minus the water mean for this simulated sample.

$$\bar{x}_B - \bar{x}_W = -0.84$$

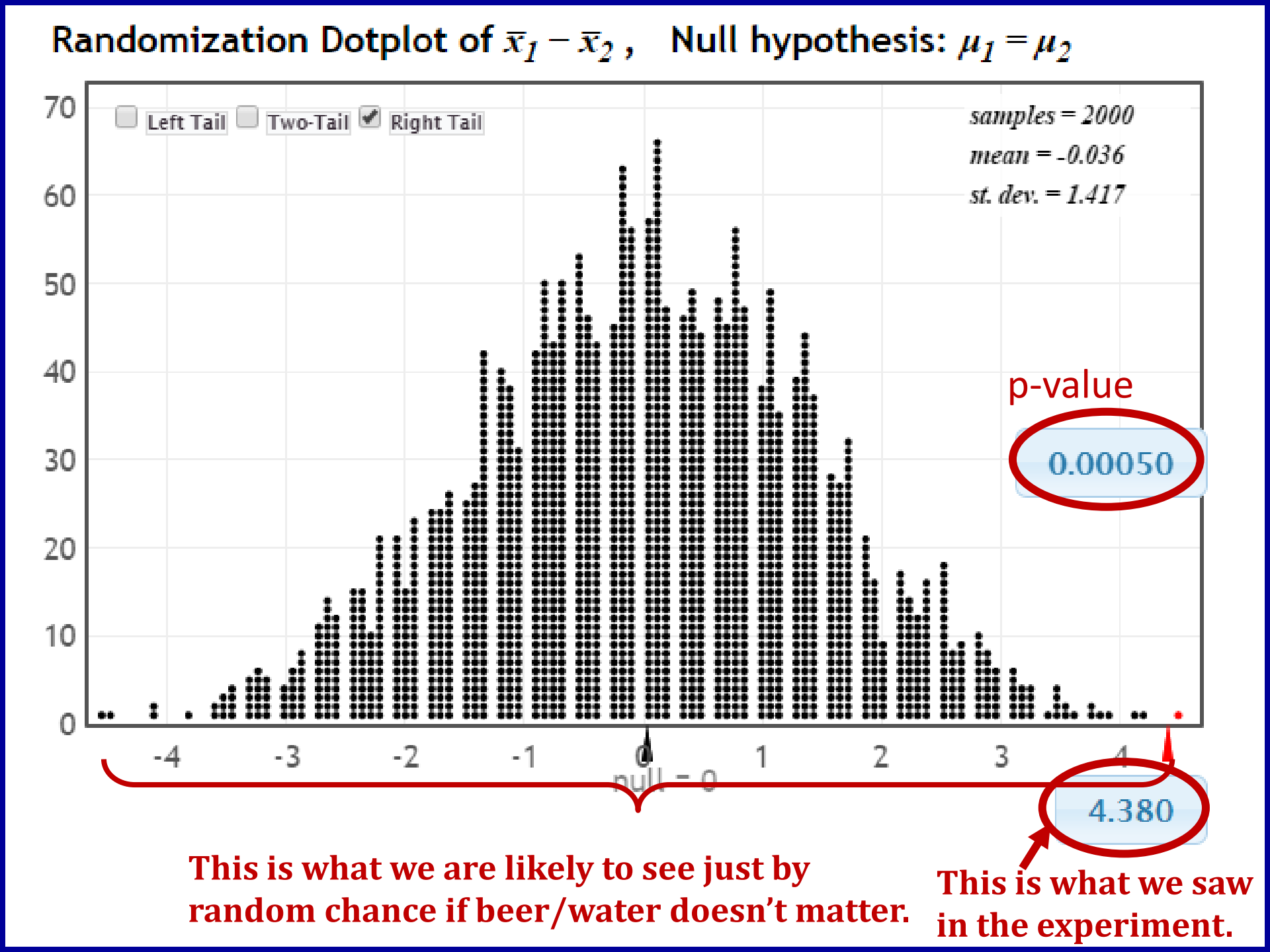
4. Do this thousands of times!

We need technology!!

STATKEY!

www.lock5stat.com/statkey

Free, online, works on all platforms, easy to use


Randomization Dotplot of $\bar{x}_1 - \bar{x}_2$, Null hypothesis: $\mu_1 = \mu_2$
Left Tail
Two-Tail
Right Tail
samples = 2000
mean = -0.036
st. dev. = 1.417
p-value
0.00050
70
60
50
40
30
20
10
0
-4
-3
-2
-1
0
1
2
3
4
null = 0
4.380
This is what we are likely to see just by random chance if beer/water doesn't matter.
This is what we saw in the experiment.

Beer and Mosquitoes
The Conclusion!

The results seen in the experiment are very unlikely to happen just by random chance (less than 1 out of 1000!)

We have strong evidence that drinking beer does attract mosquitoes!

Another Look at Beer/Mosquitoes

1. Check conditions
2. Which formula?
5. Which theoretical distribution?

$$t = \frac{\bar{x}_B - \bar{x}_W}{\sqrt{\frac{s_B^2}{n_B} + \frac{s_W^2}{n_W}}}$$

6. df?
7. Find p-value
8. Interpret a decision

3. Calculate numbers and plug into formula

$$t = \frac{23.6 - 19.22}{\sqrt{\frac{4.1^2}{25} + \frac{3.7^2}{18}}}$$

4. Chug with calculator

$t = 3.68$

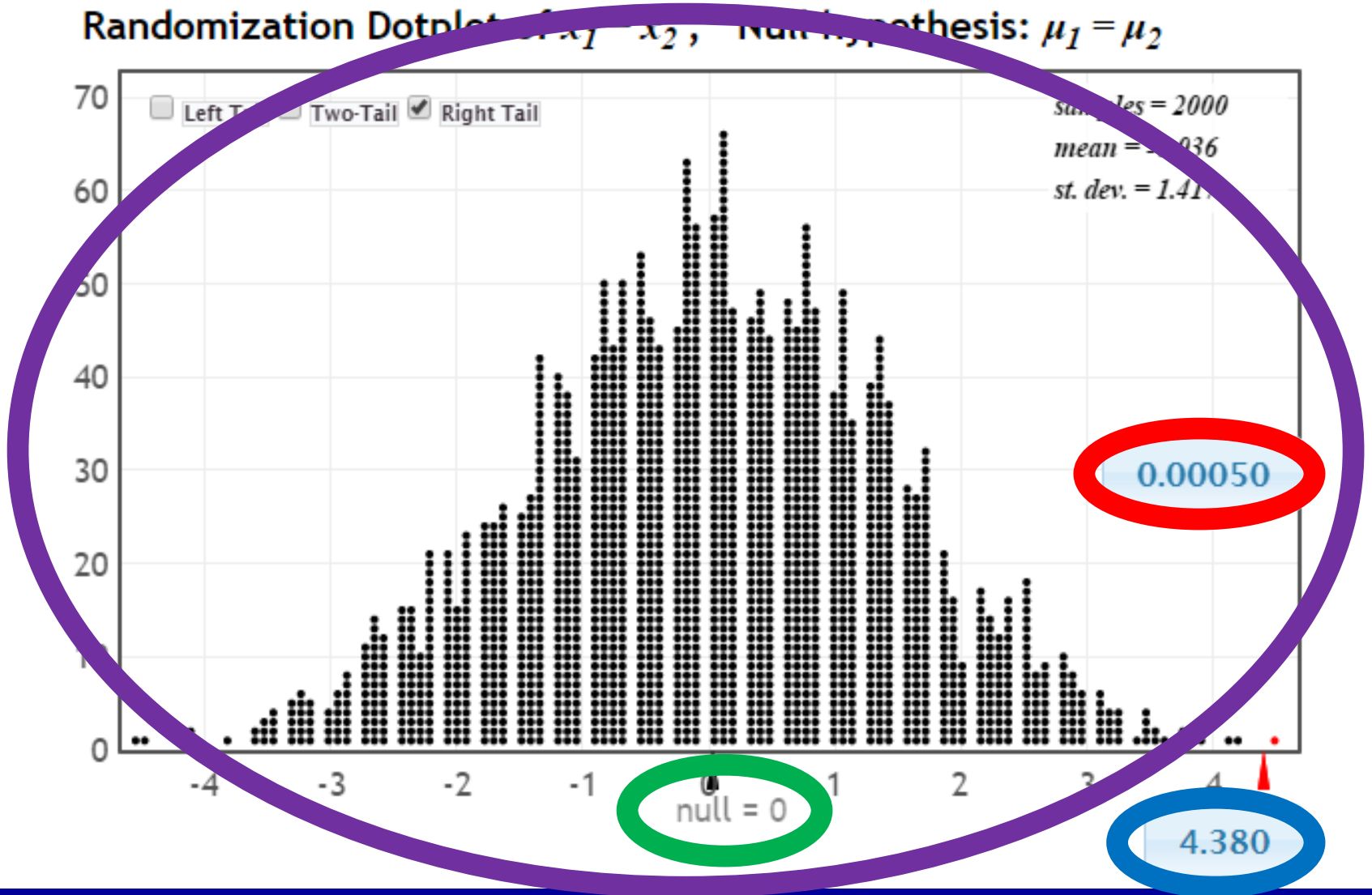
0.0005 < p-value < 0.001

TABLE B: *t*-DISTRIBUTION CRITICAL VALUES

	Tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.002	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

What's a p-value?!?

p-value: The chance of obtaining a statistic as extreme as that observed, just by random chance, if the null hypothesis is true



Conclusions are the same, but the process is very different!

"Students' approach to p-values … was procedural … and [they] did not attach much meaning to p-values"

-- Aquilonius and Brenner, "Students' Reasoning about P-Values", SERJ, November 2015

Simulation-Based Inference

- Visual!
- Intuitive!
- Easily incorporates active learning!
- Ties directly to key concepts!
- Same process for all parameters!
- More generalizable!
- No theoretical distributions! No formulas! No formal probability!
- No algebra!!!

Intro Stats

Why is there a push now to change?

Technology!!!

We now have the technological power to do thousands of simulations quickly and easily.

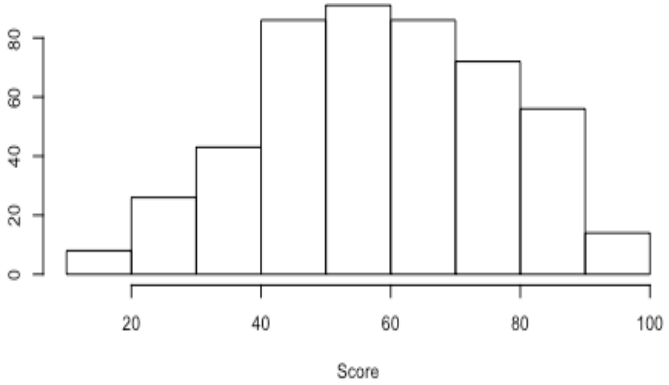
Ways to Change

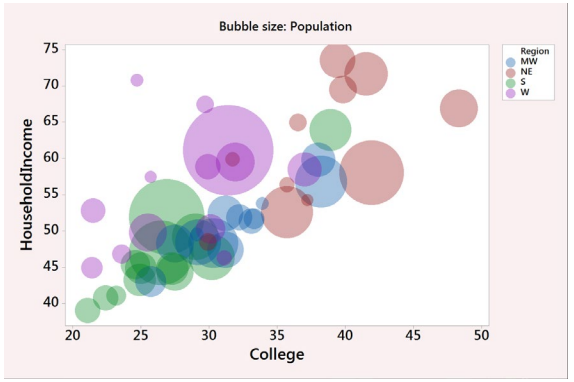
1. Make everything about real **data**
2. Reduce tedious "by-hand" work and let **technology** do the heavy-lifting
3. Focus on interpretation and **concepts**
4. Improve conceptual understanding and reduce reliance on prerequisites with **simulation-based inference**
5. Embrace the power of **data visualization** and multivariable thinking

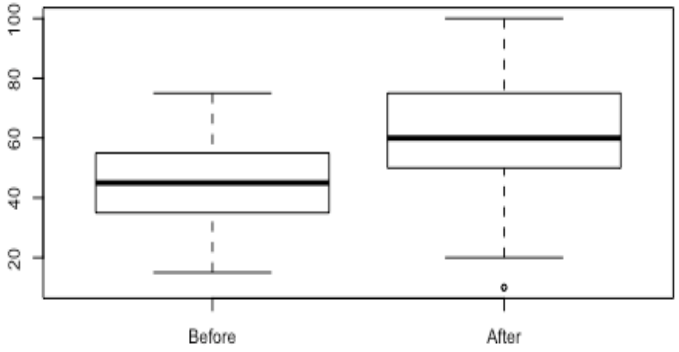
2016 GAISE Guidelines

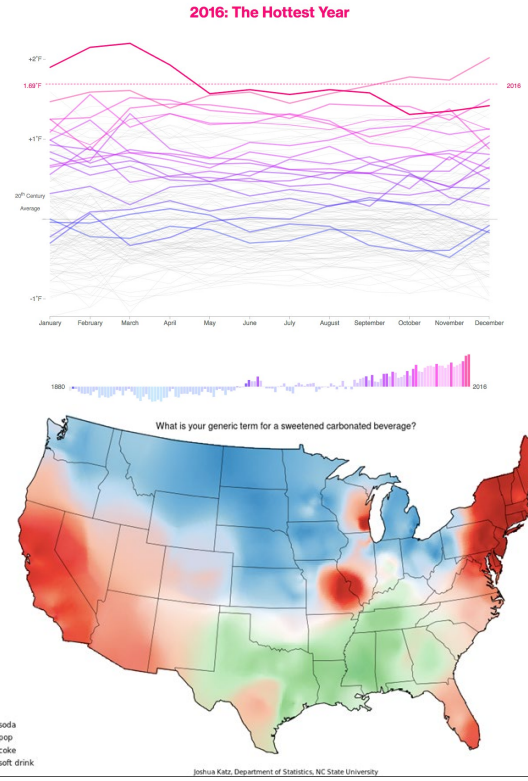
1. Teach statistical thinking.
   - Teach statistics as an investigative process of problem-solving and decision making.
   - Give students experience with multivariable thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning.

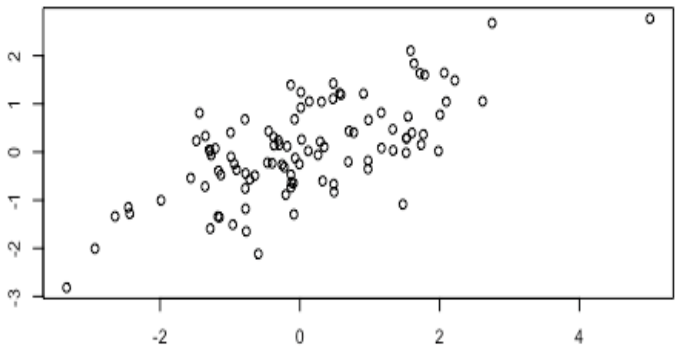
Data Visualization



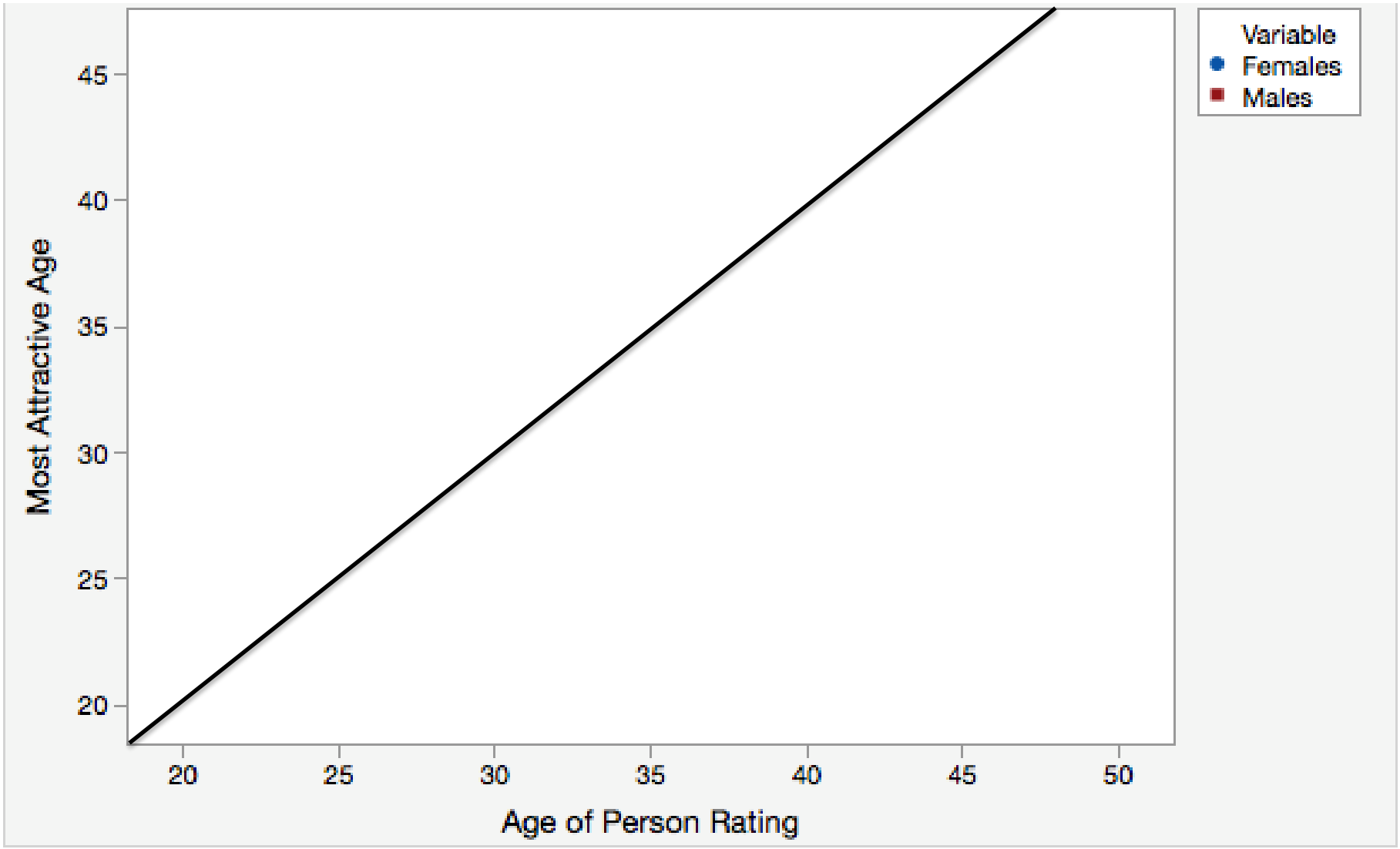




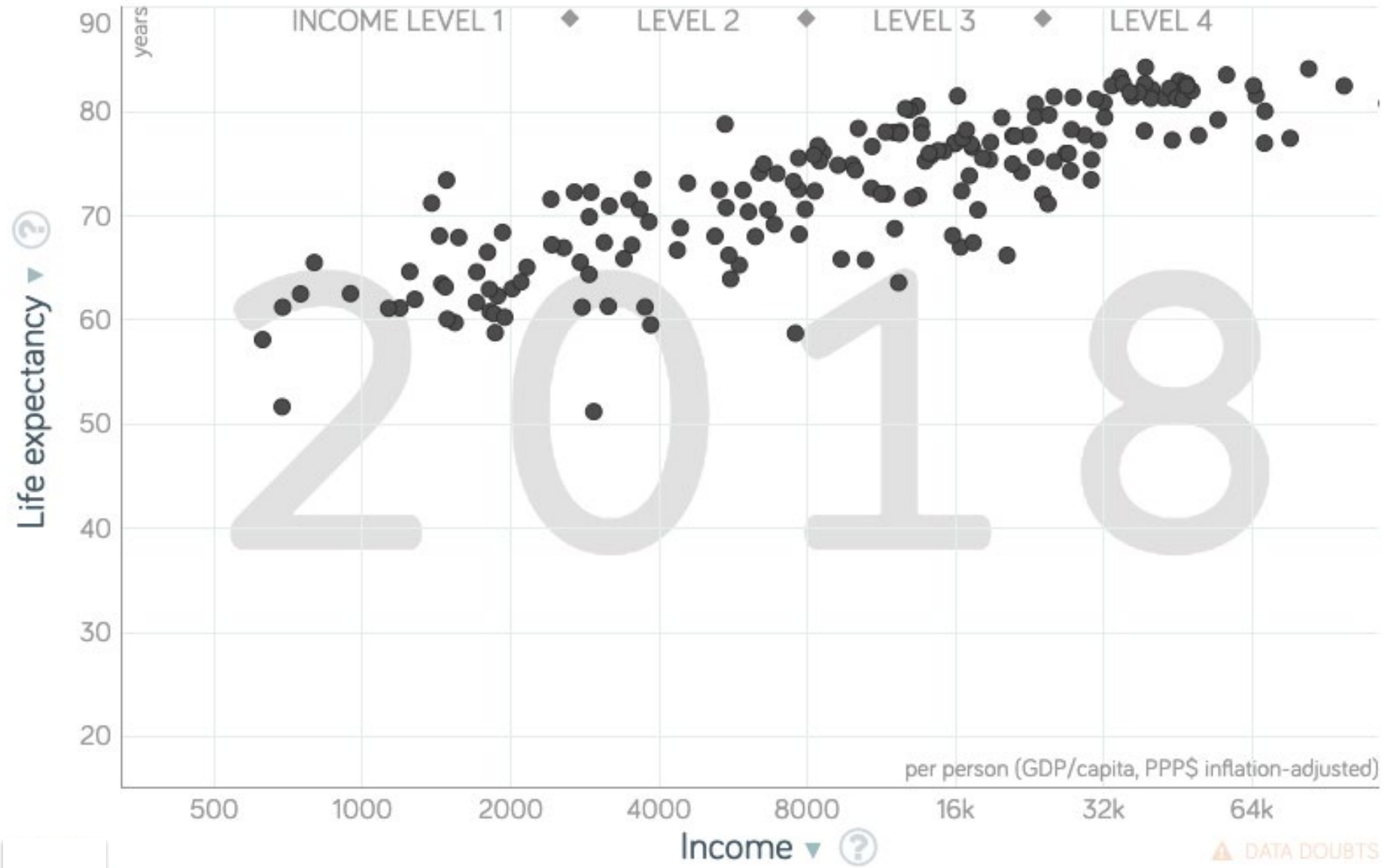




OK Cupid Data

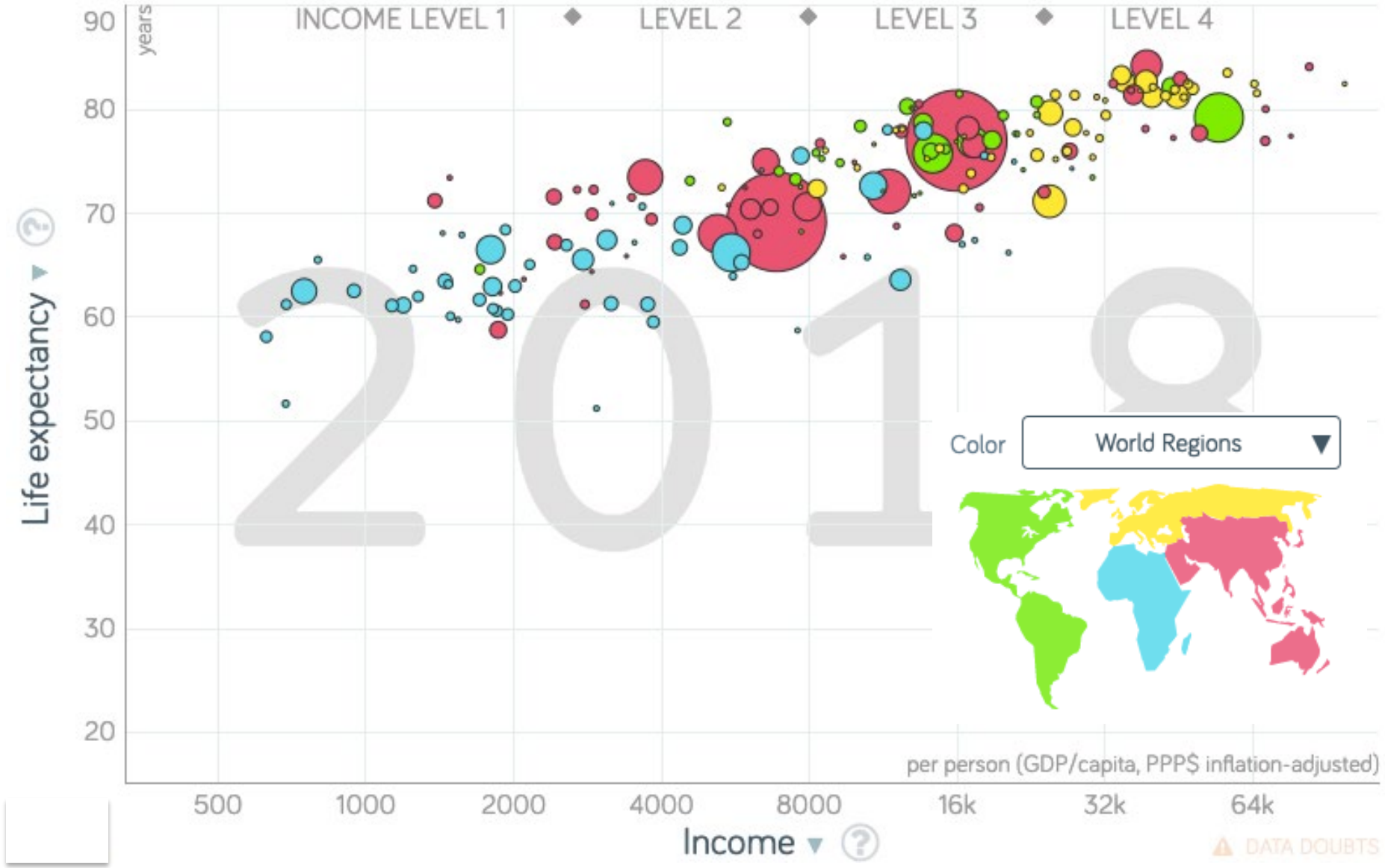


Scatterplot



Scatterplot: Add a 3rd Variable!

Scatterplot: Add a 4th Variable!



But we want more!

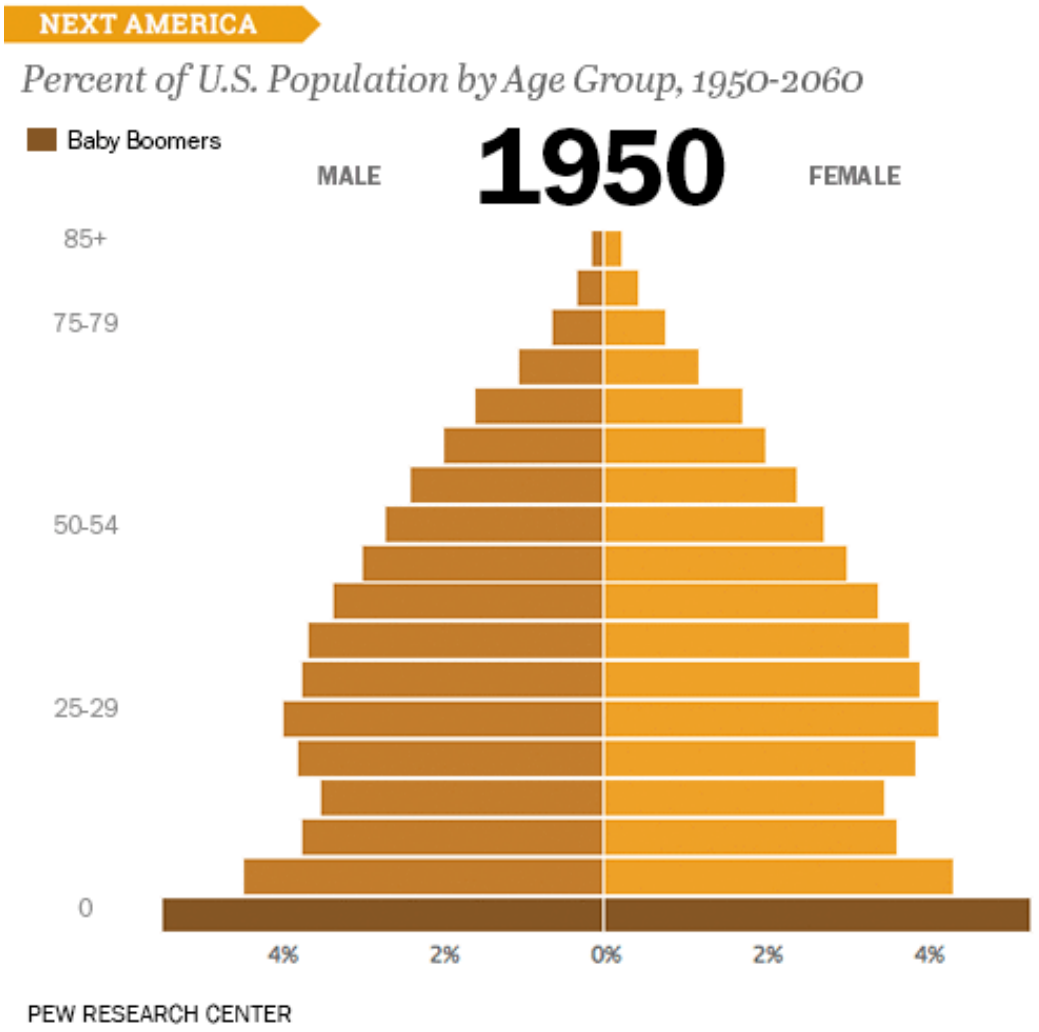
Make it INTERACTIVE!

Make it DYNAMIC!

www.gapminder.com/tools

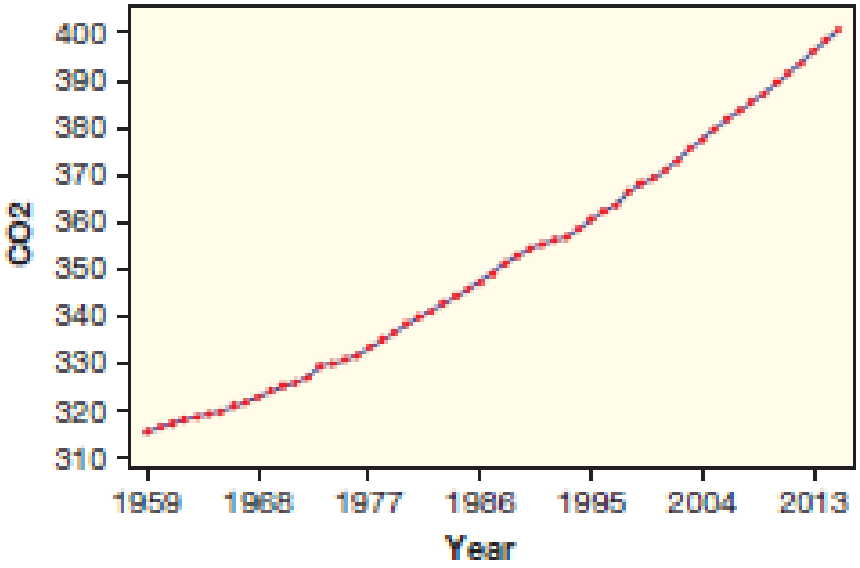
Hans Rosling's 200 Countries, 200 Years, 4 Minutes
The Joy of Stats – BBC Four

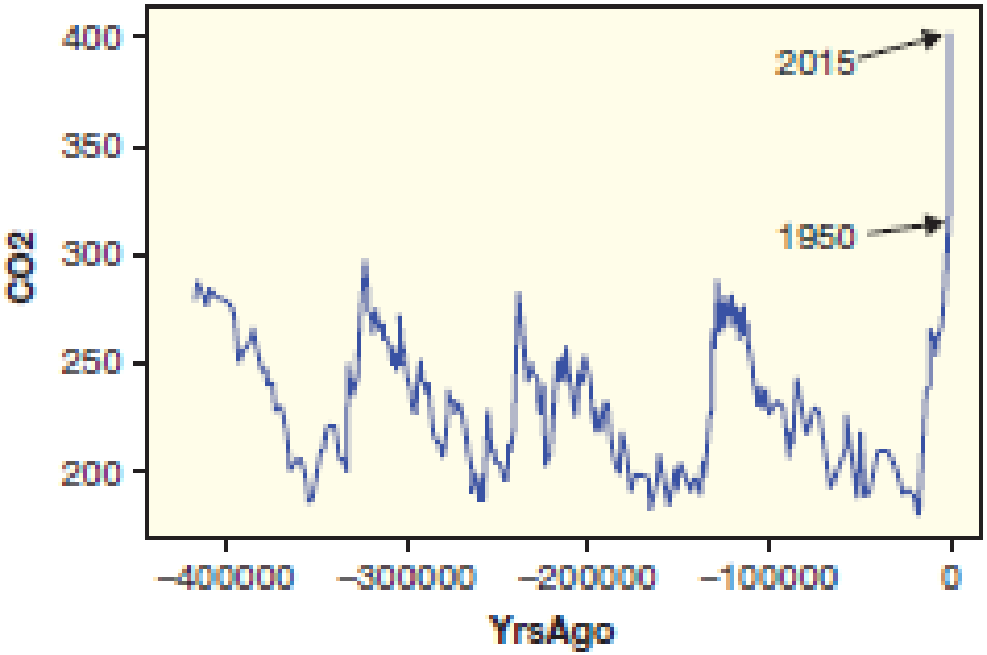
US Population by Age and Sex



www.pewresearch.org/next-america/#Two-Dramas-in-Slow-Motion

Data Over Time: CO_2 Levels

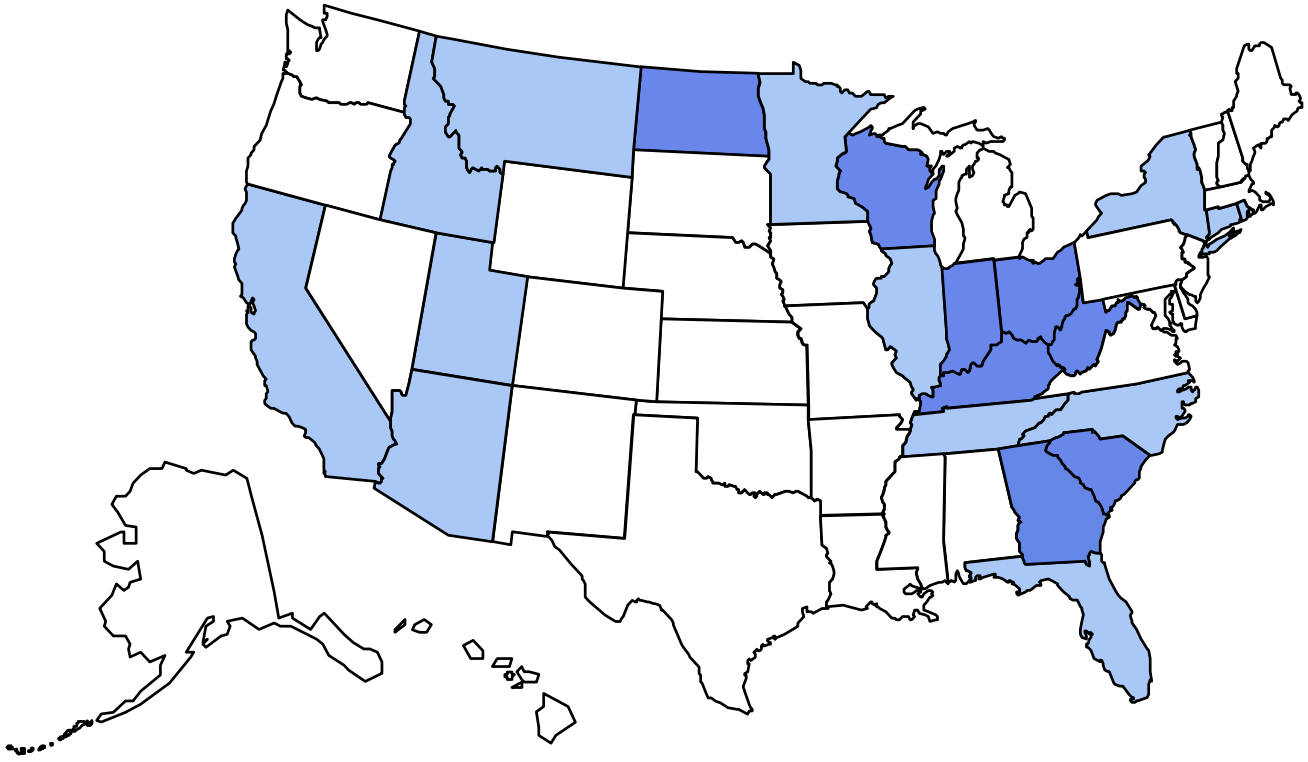




A whole class of global warming data visualizations

Obesity* in U.S. Adults: 1985

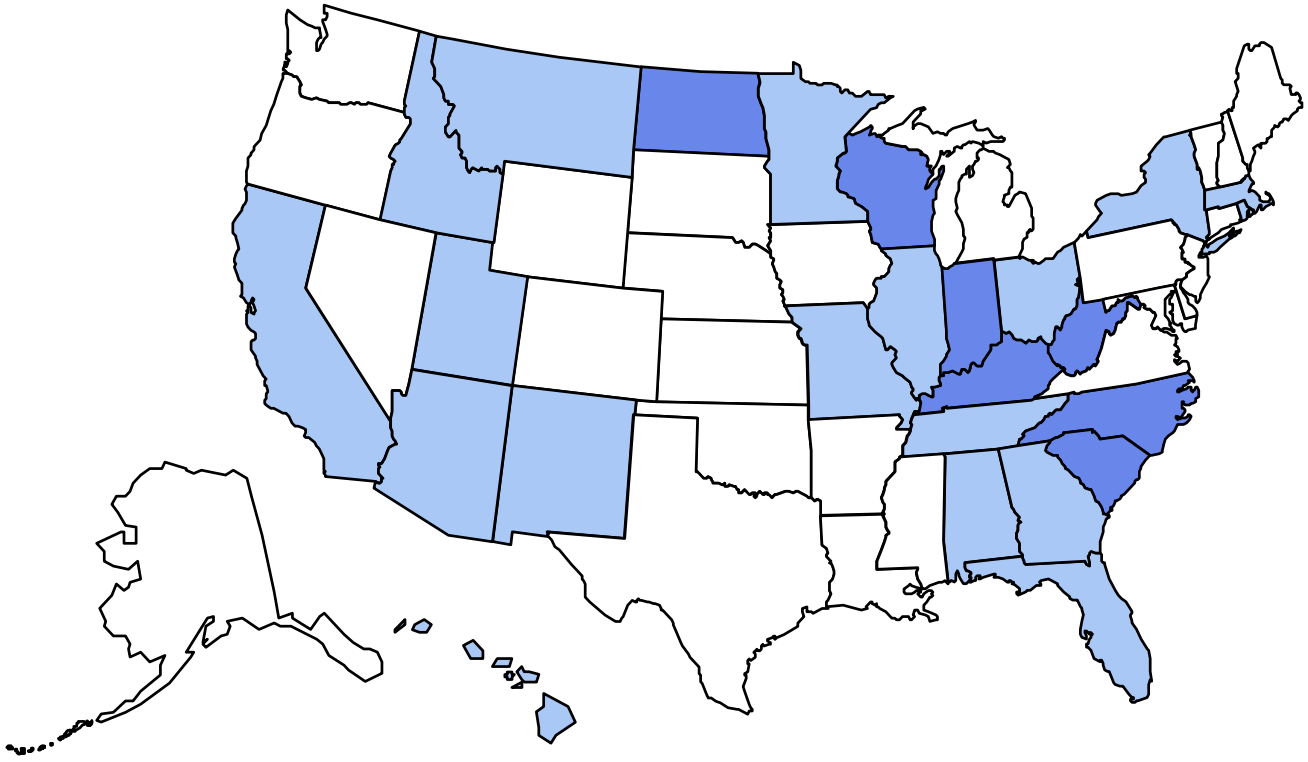
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

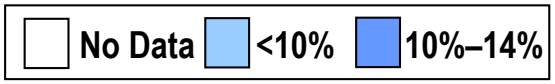


No Data | <10% | 10%–14%

Obesity* in U.S. Adults: 1986

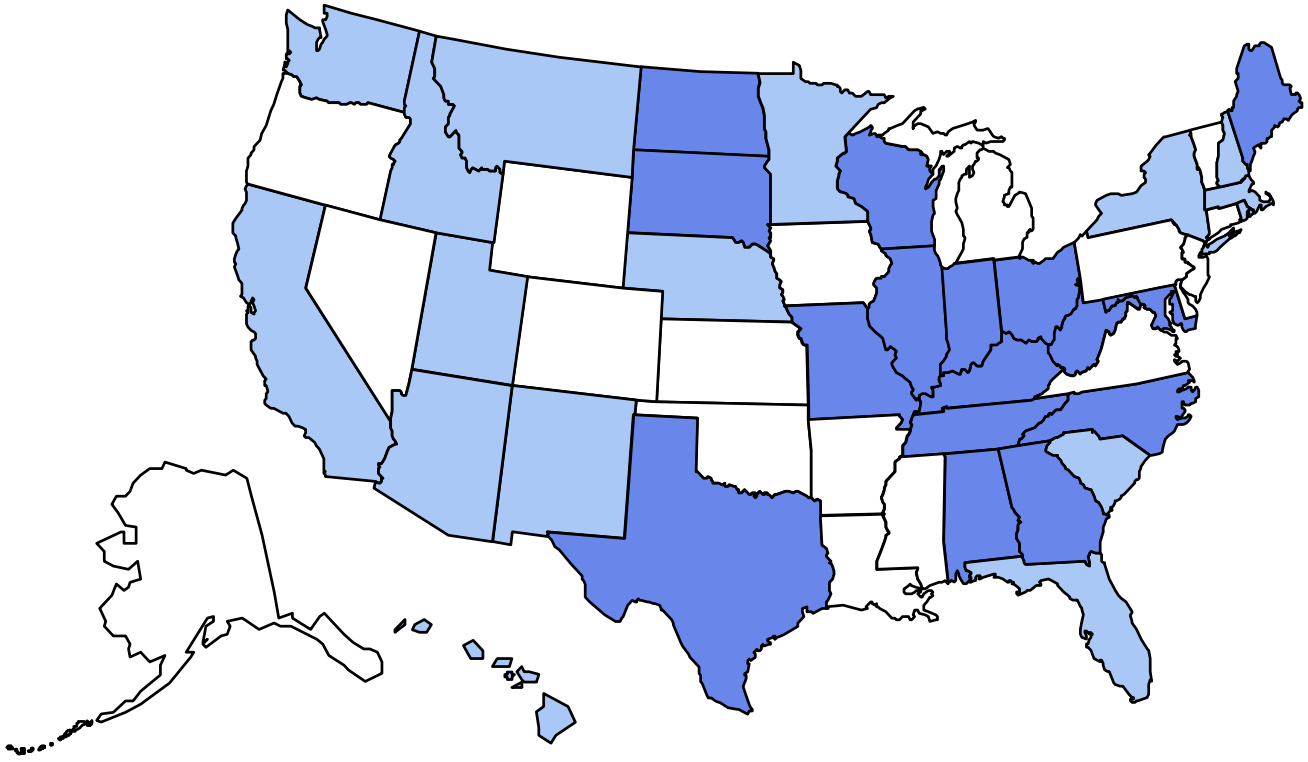
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)





Obesity* in U.S. Adults: 1987

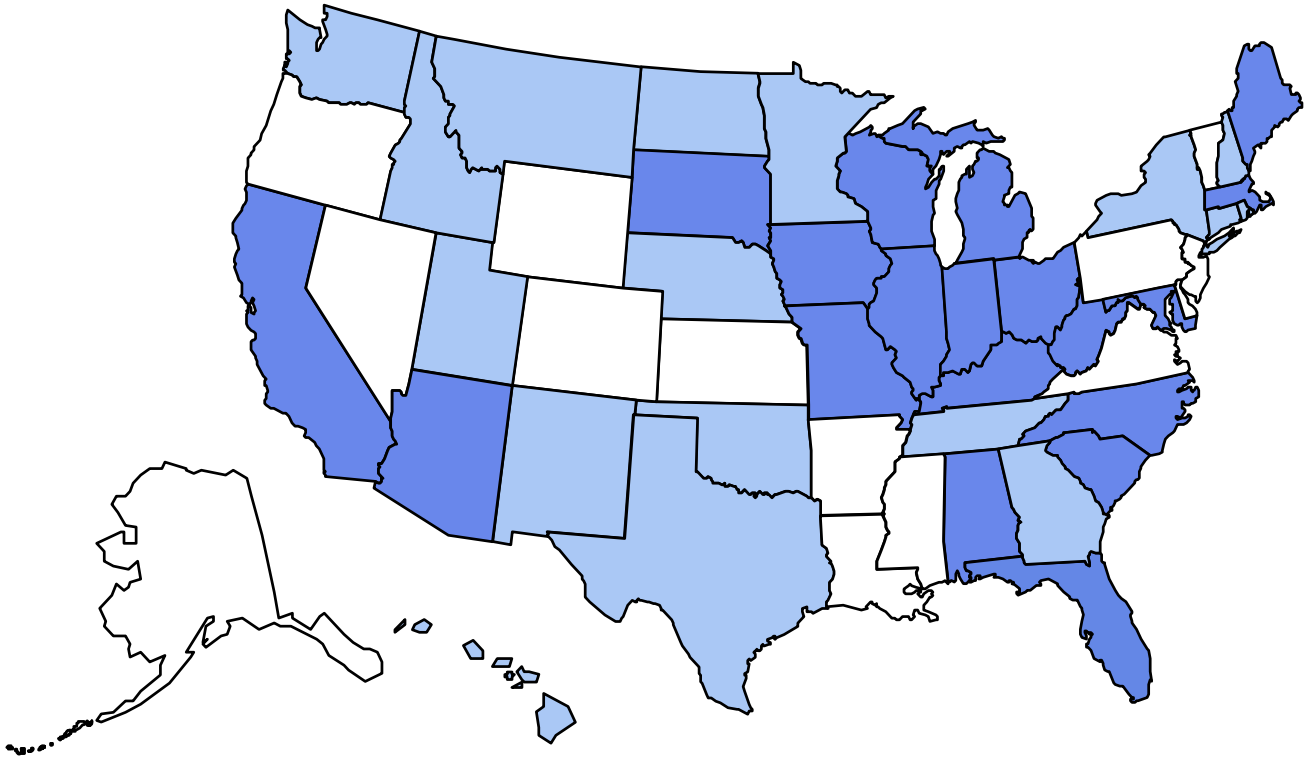
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

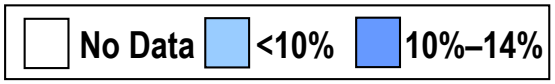


No Data | <10% | 10%–14%

Obesity* in U.S. Adults: 1988

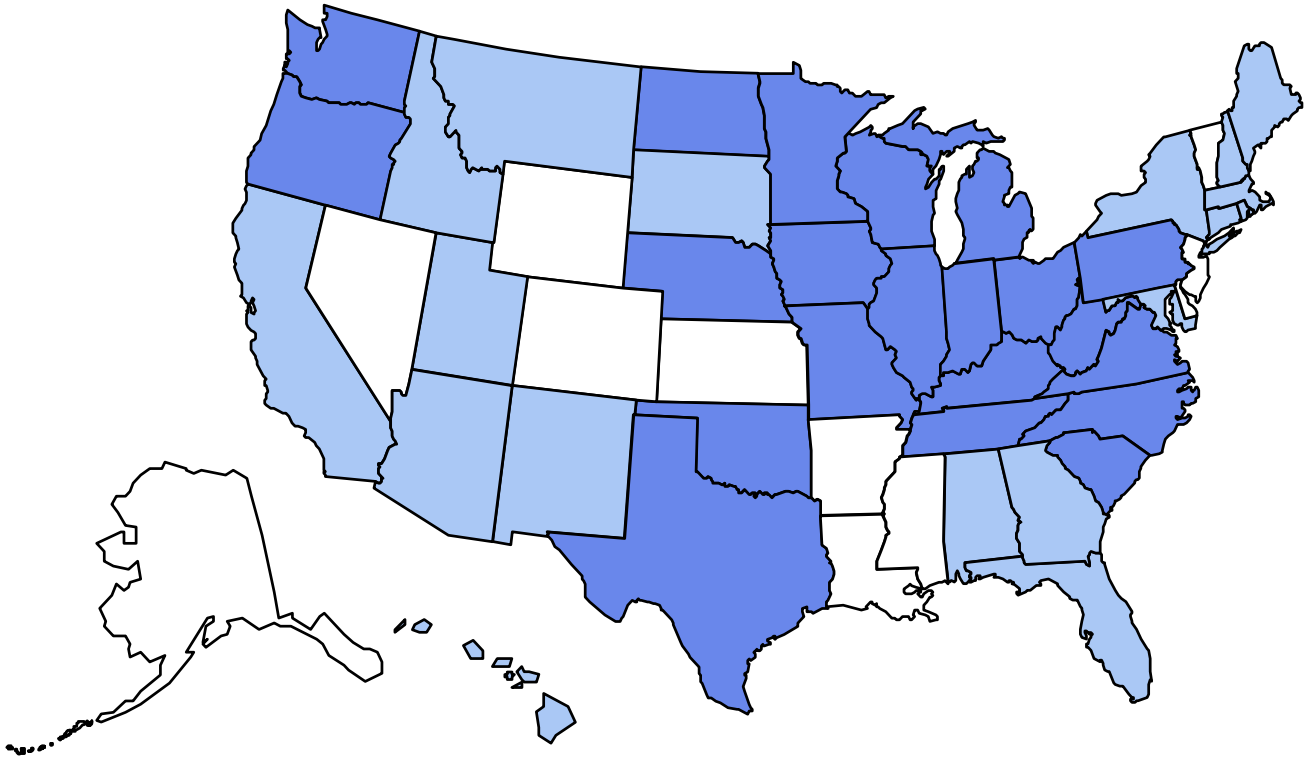
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)





Obesity* in U.S. Adults: 1989

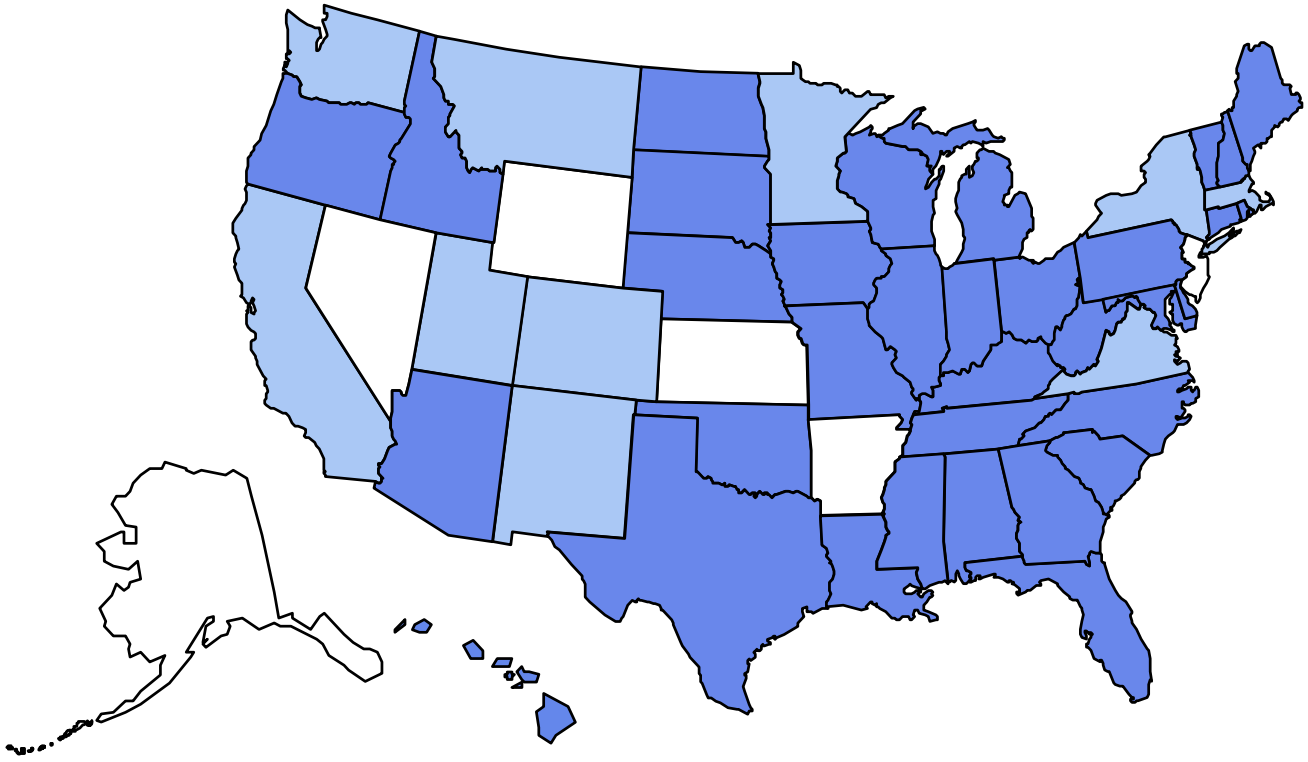
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



No Data | <10% | 10%–14%

Obesity* in U.S. Adults: 1990

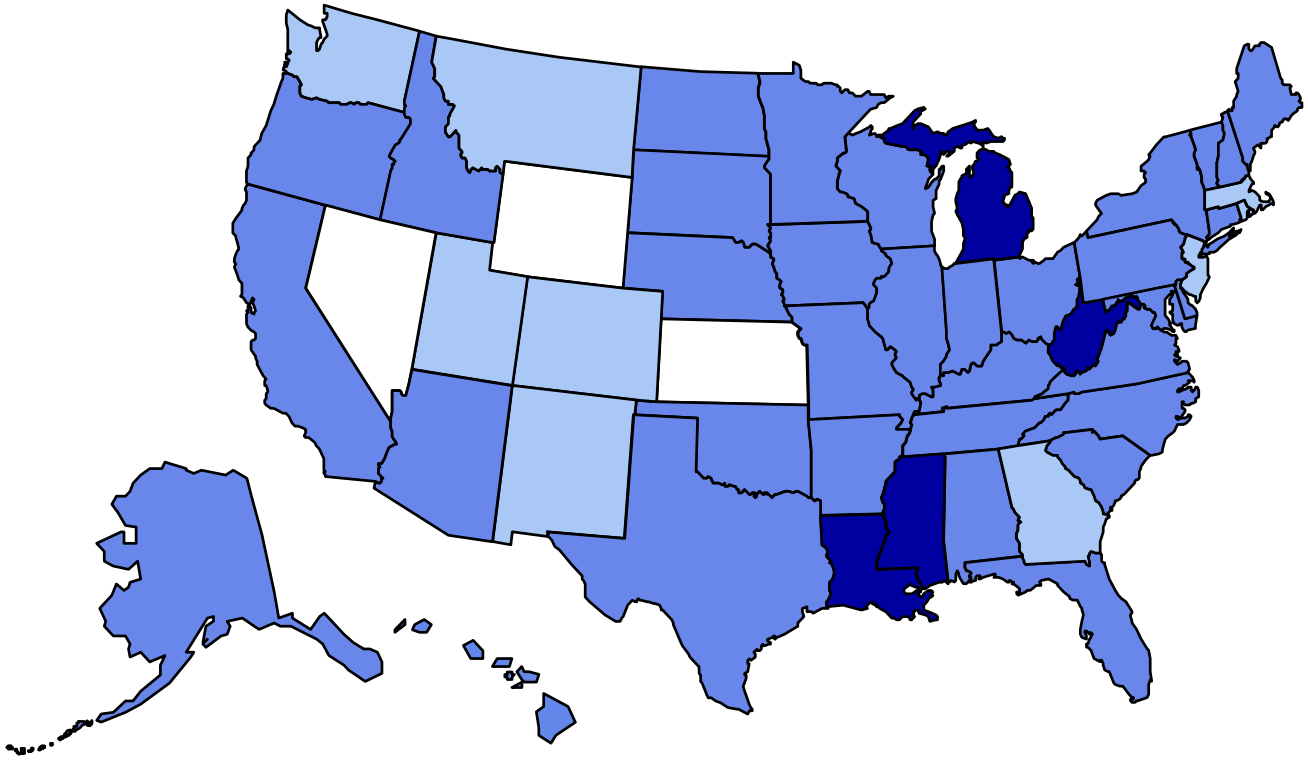
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

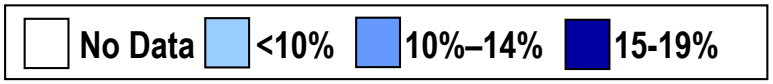


No Data <10% 10%–14%

Obesity* in U.S. Adults: 1991

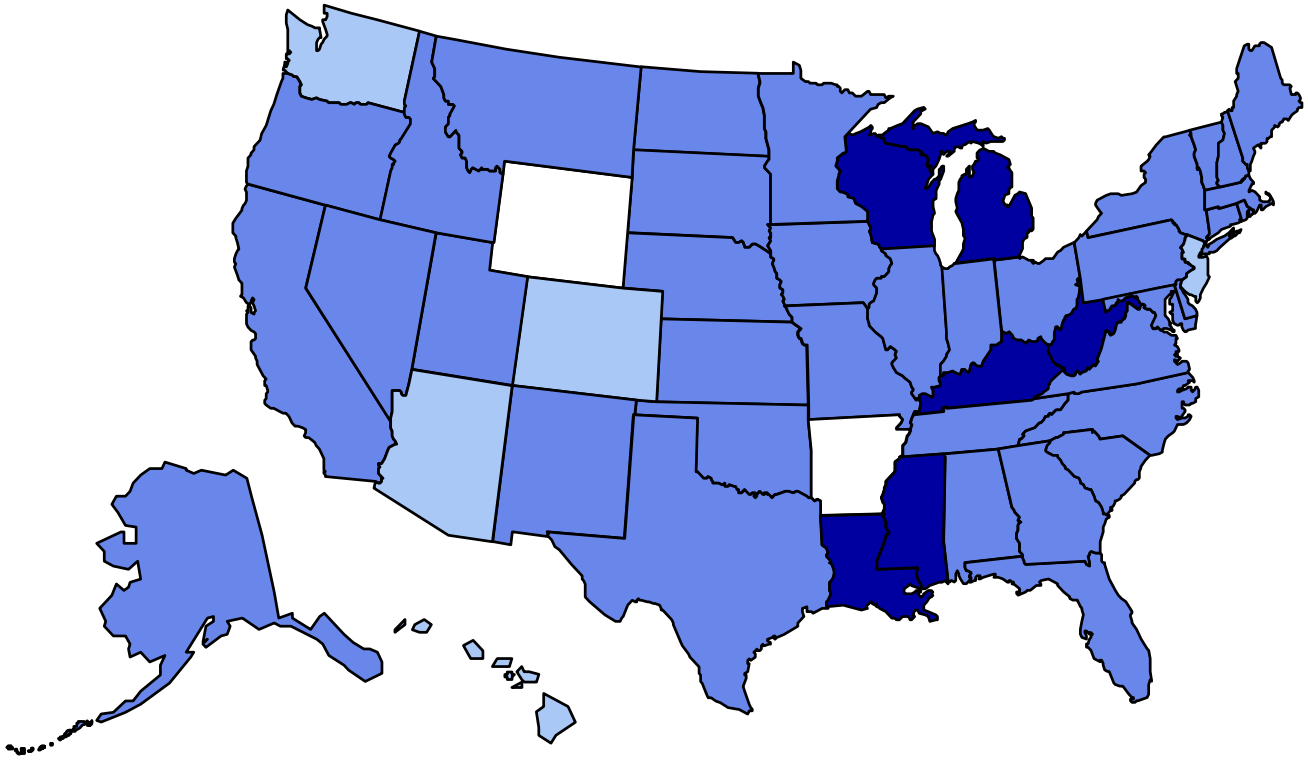
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

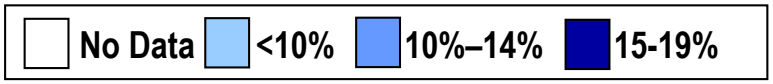




Obesity* in U.S. Adults: 1992

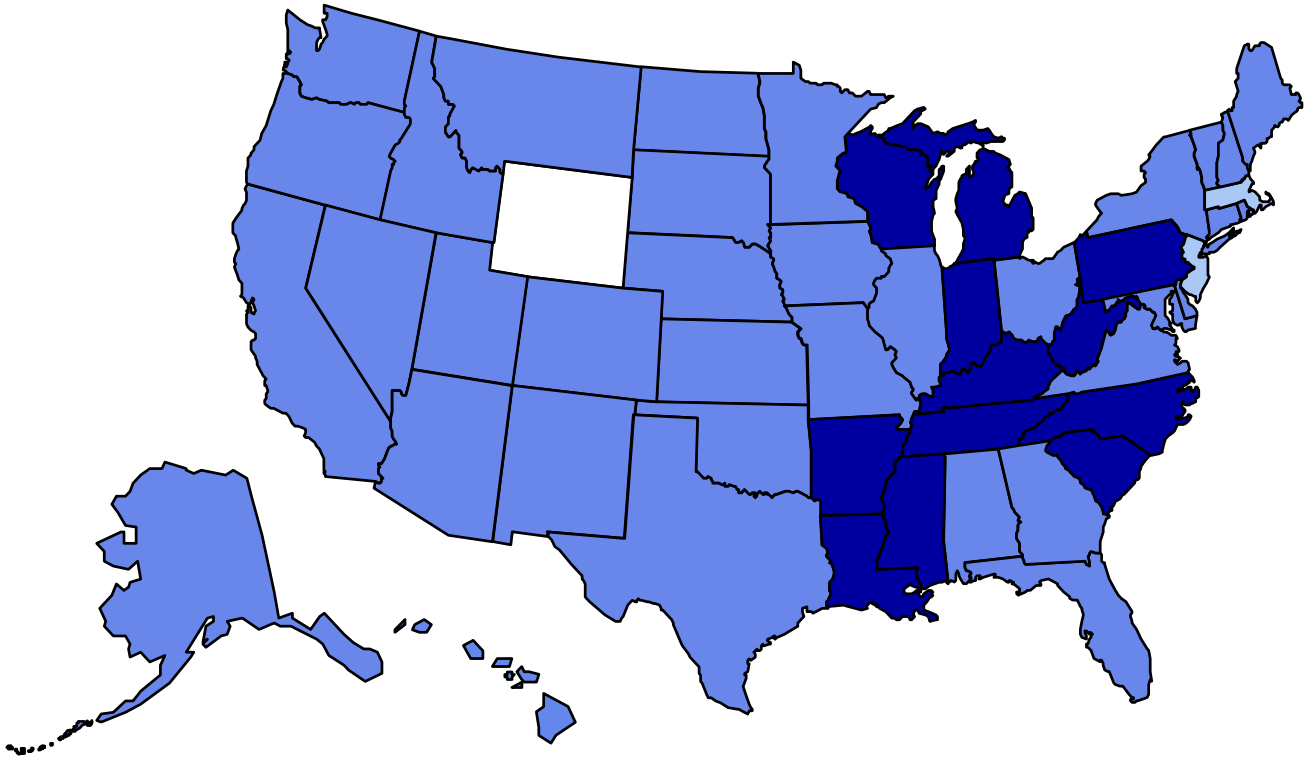
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)





Obesity* in U.S. Adults: 1993

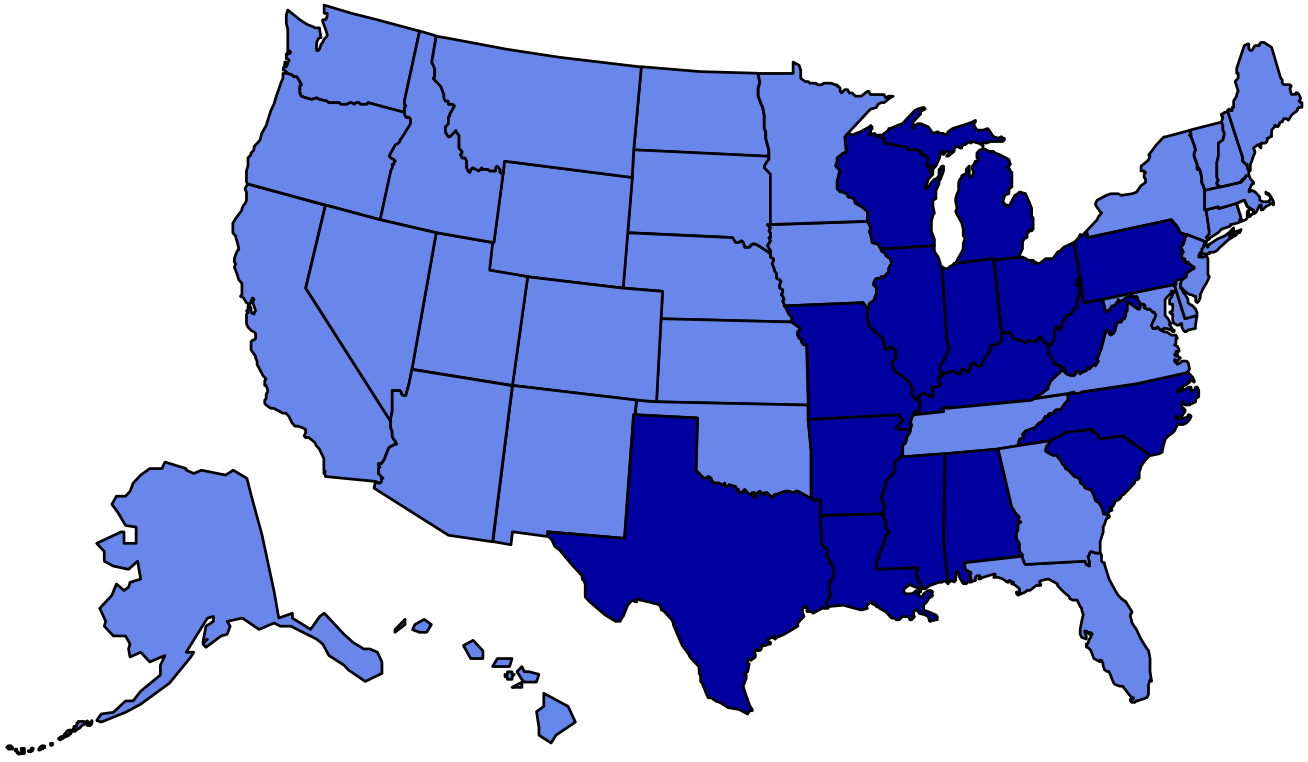
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

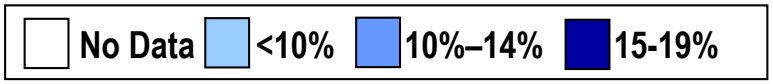


No Data | <10% | 10%–14% | 15-19%

Obesity* in U.S. Adults: 1994

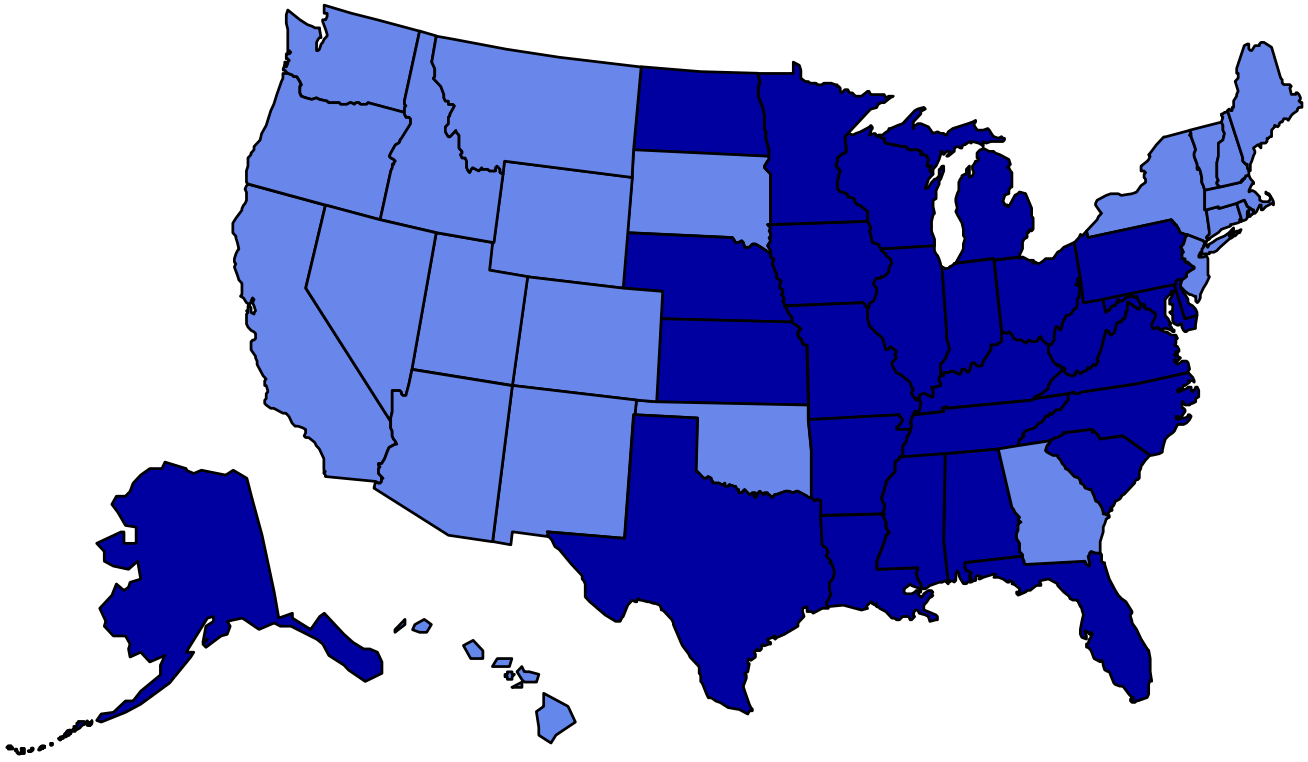
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

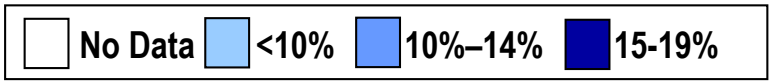




Obesity* in U.S. Adults: 1995

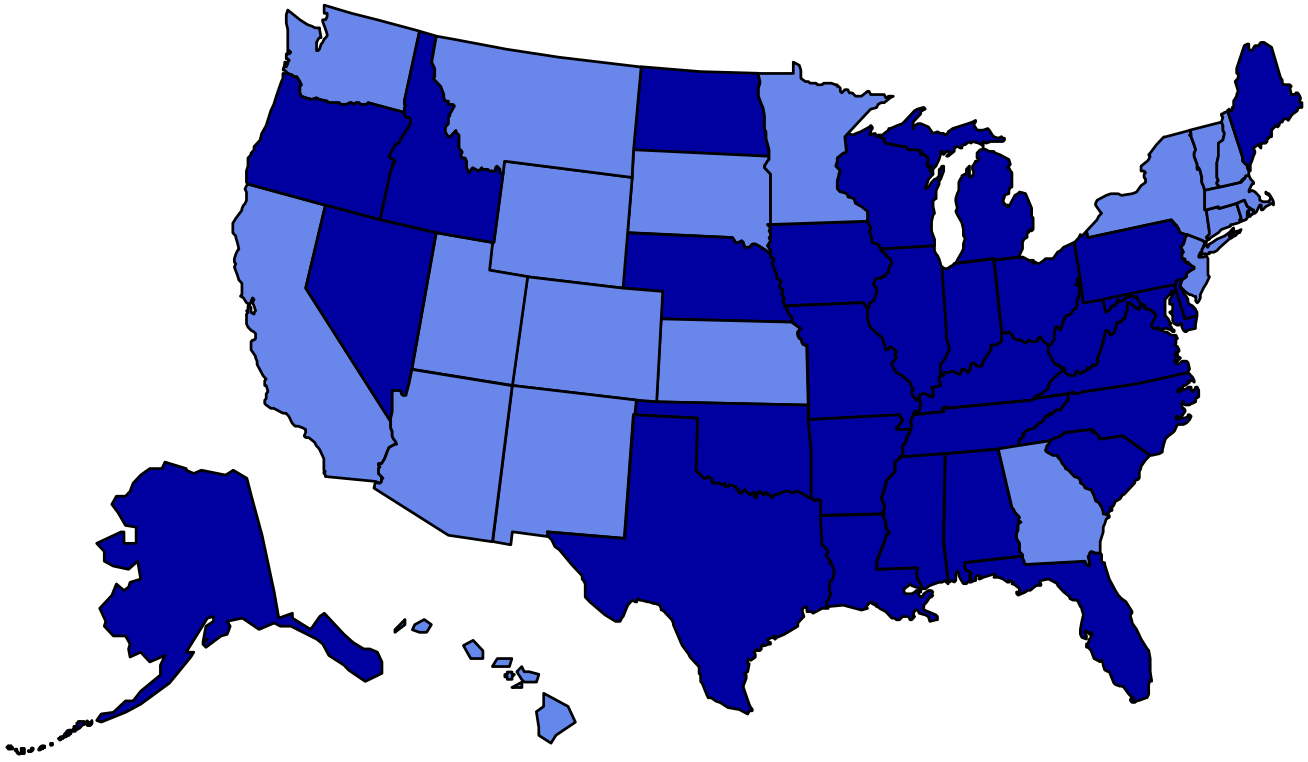
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

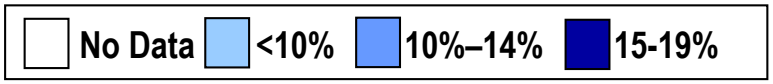




Obesity* in U.S. Adults: 1996

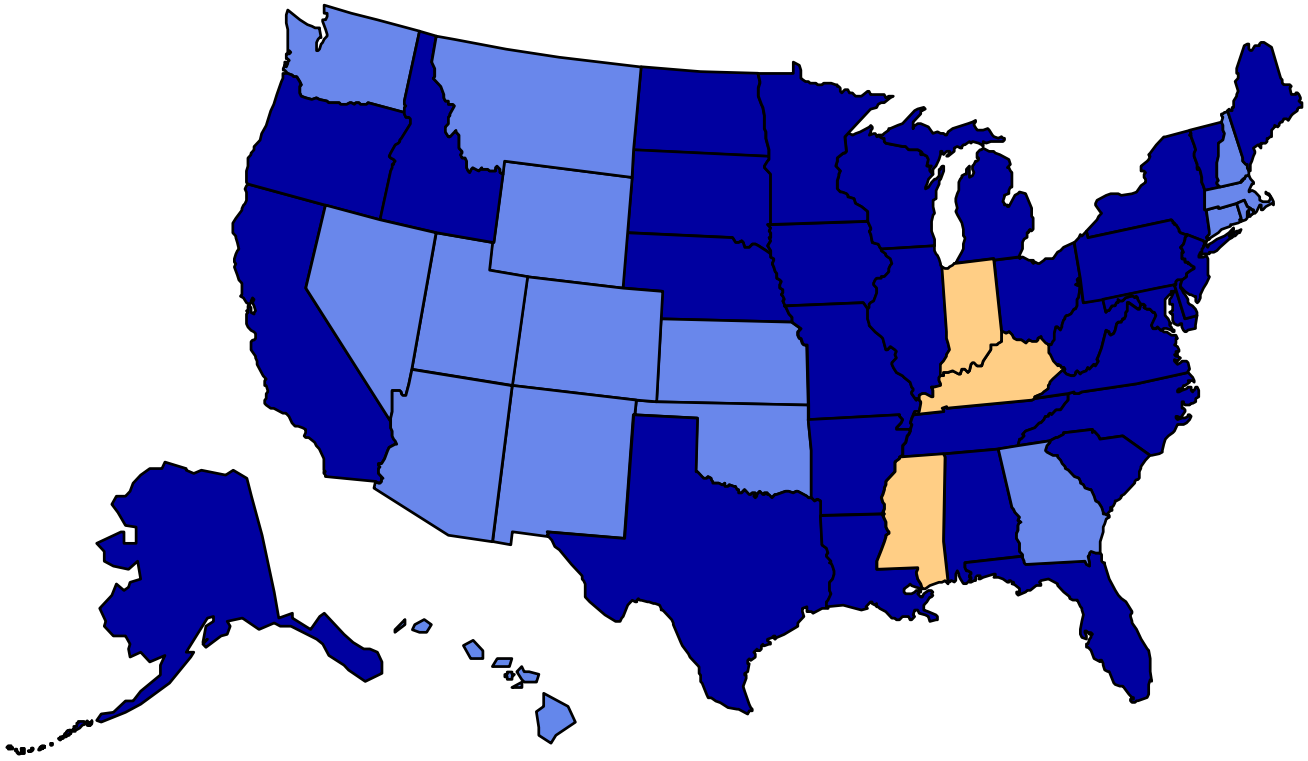
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

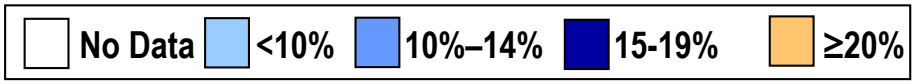




Obesity* in U.S. Adults: 1997

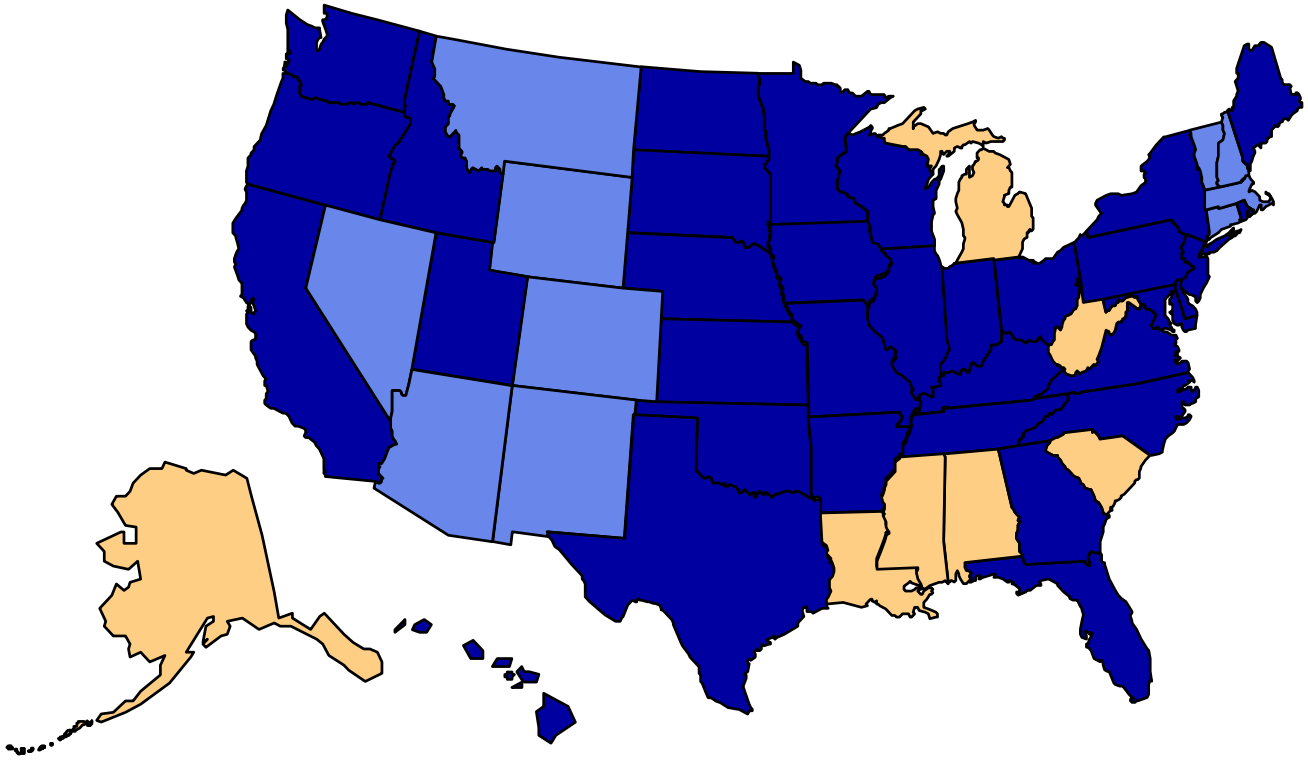
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

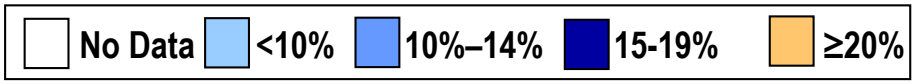




Obesity* in U.S. Adults: 1998

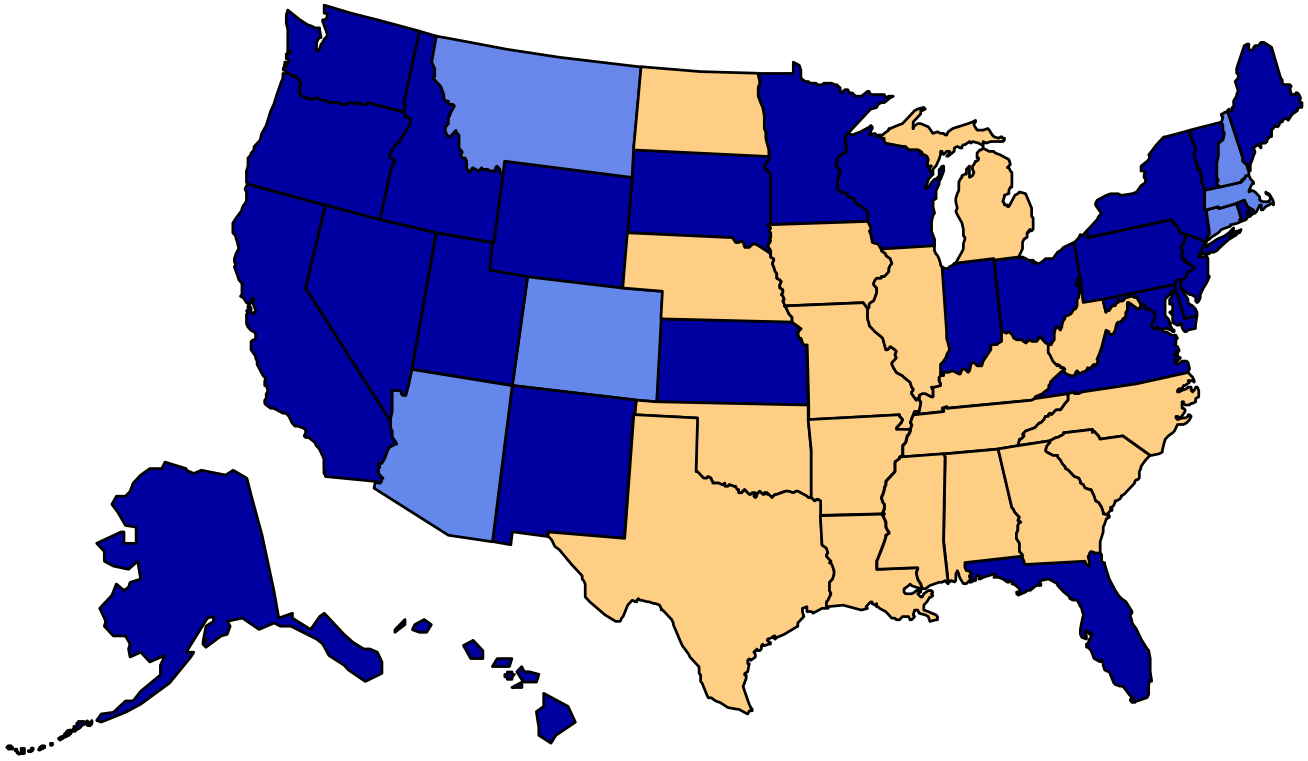
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

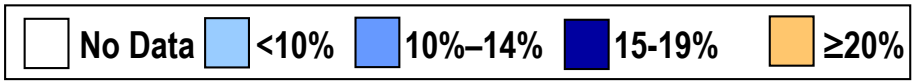




Obesity* in U.S. Adults: 1999

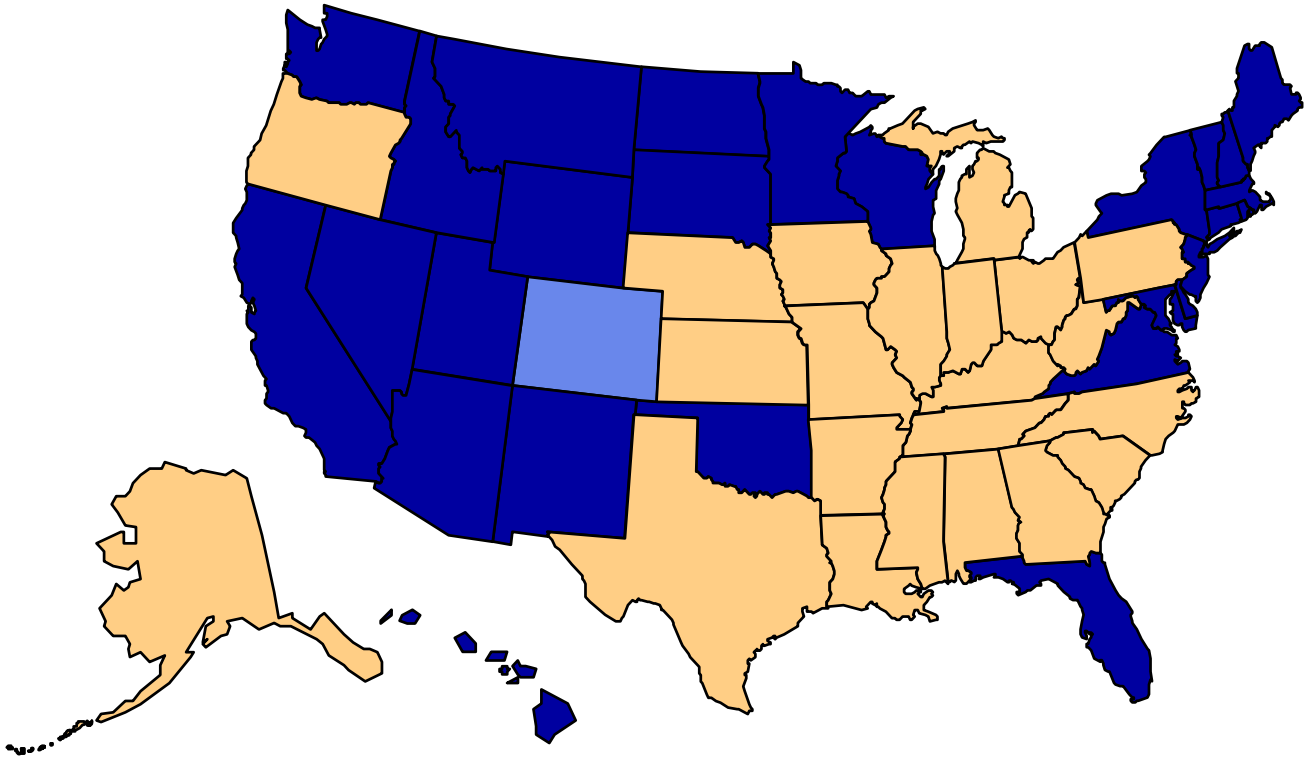
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

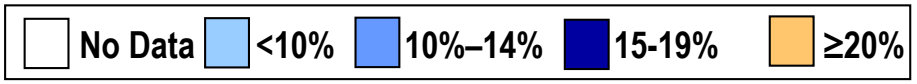




Obesity* in U.S. Adults: 2000

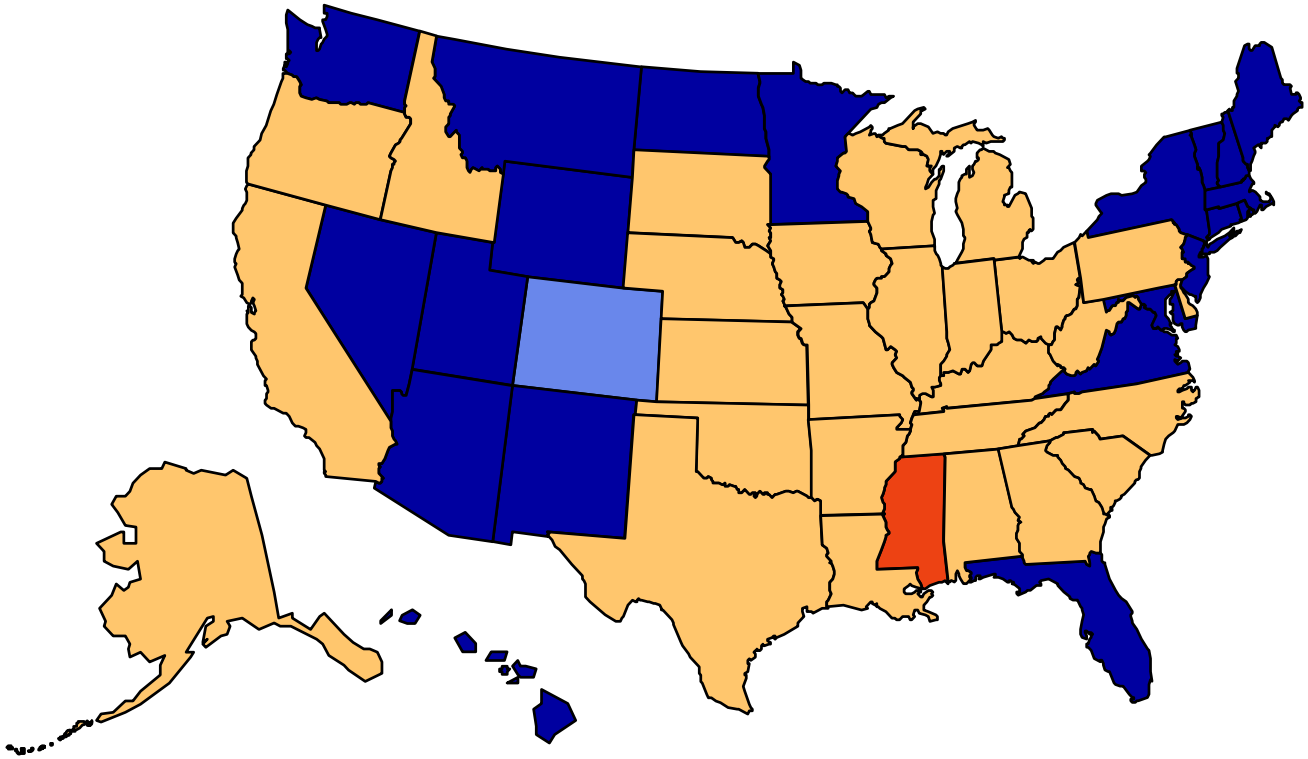
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

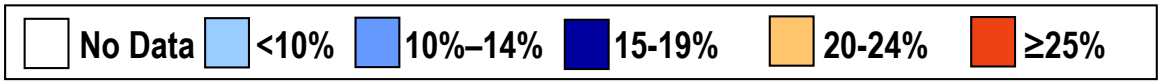




Obesity* in U.S. Adults: 2001

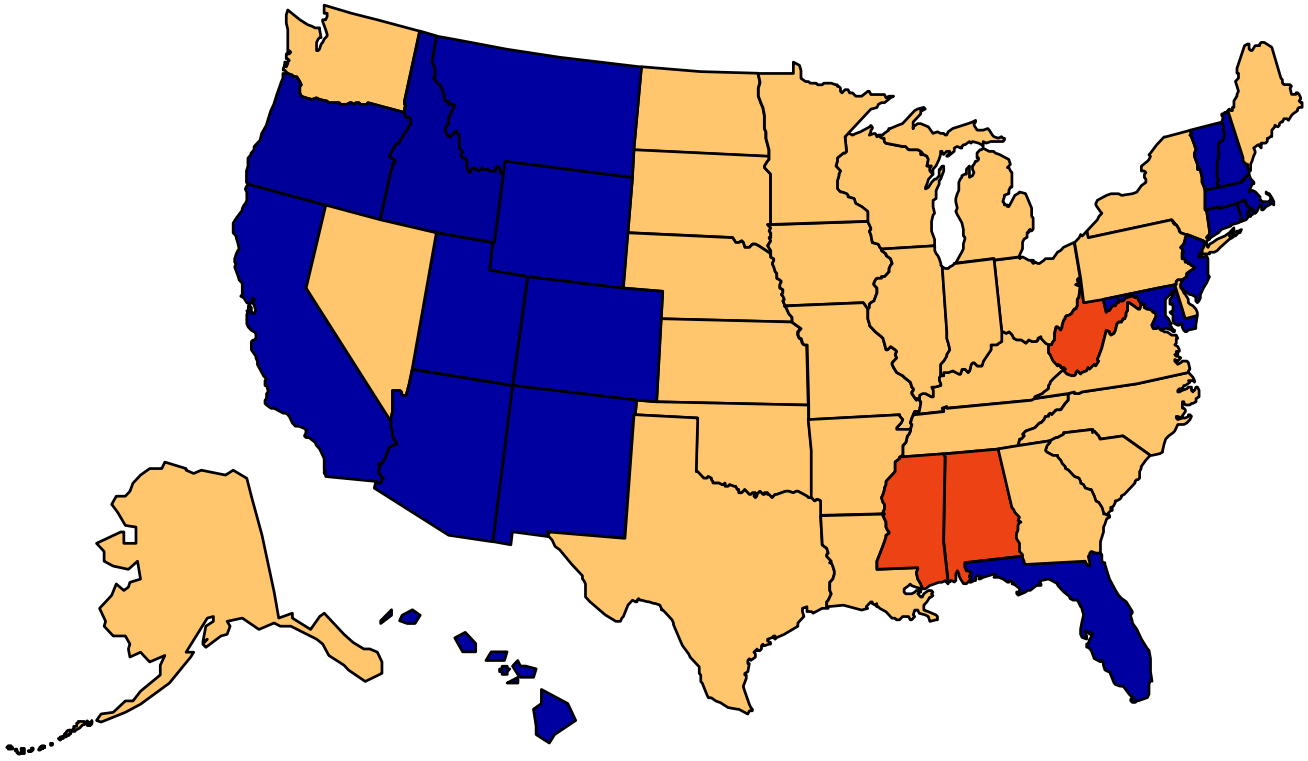
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

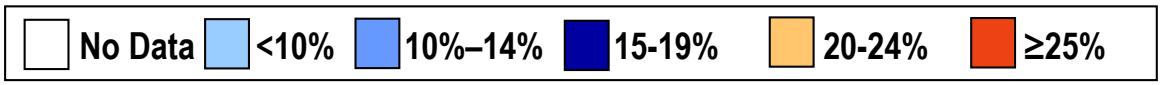




Obesity* in U.S. Adults: 2002

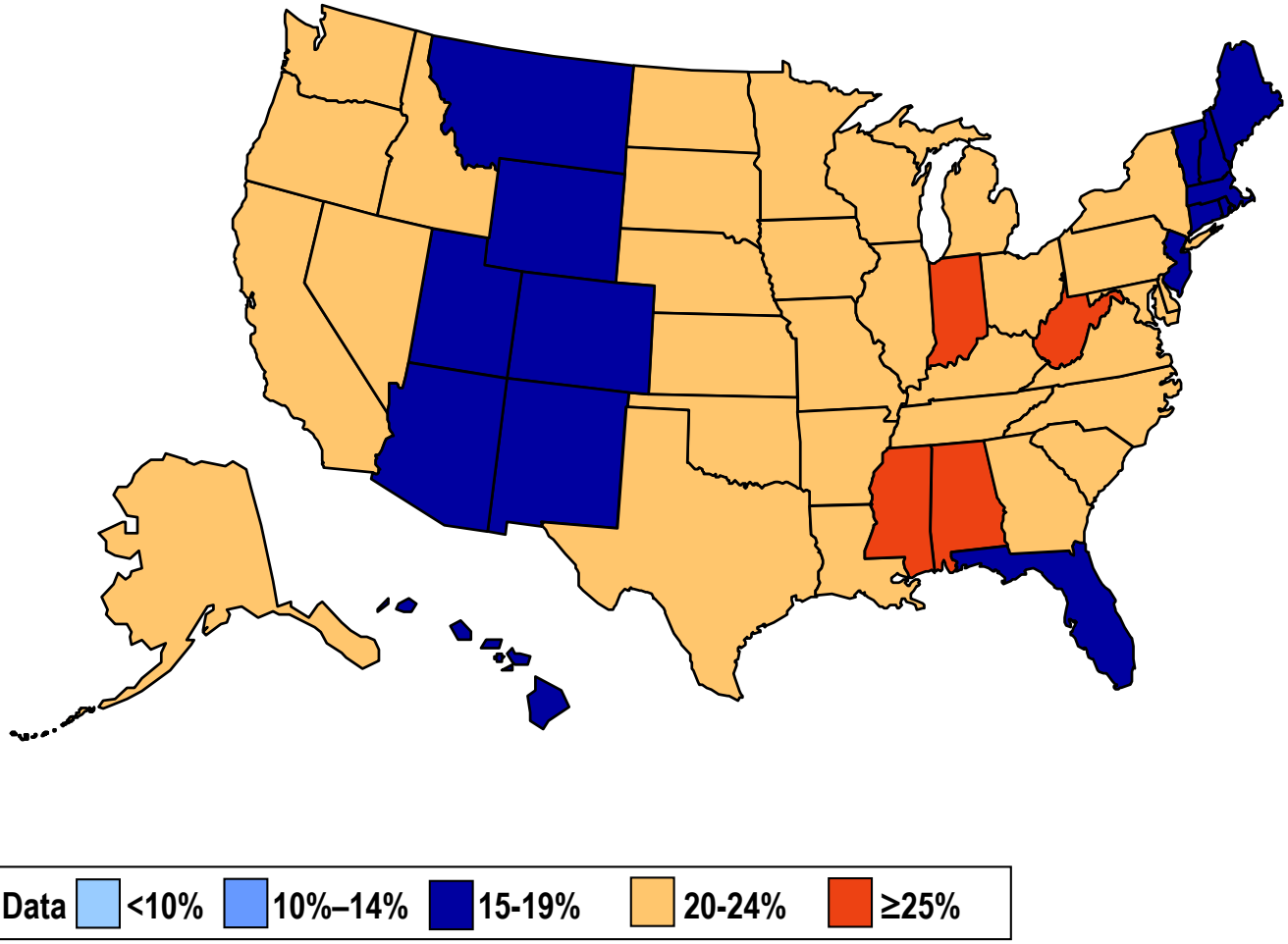
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)





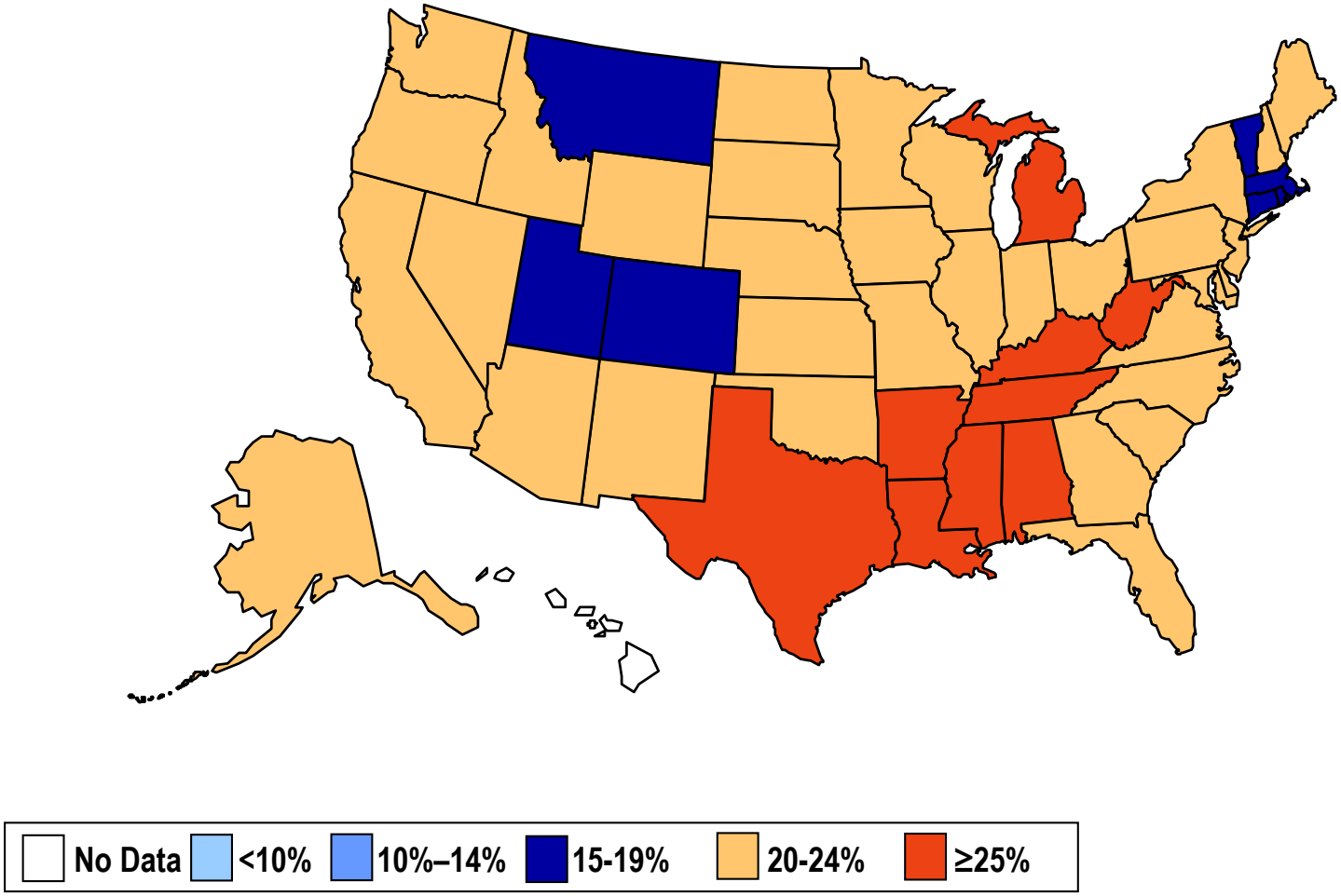
Obesity* in U.S. Adults: 2003

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



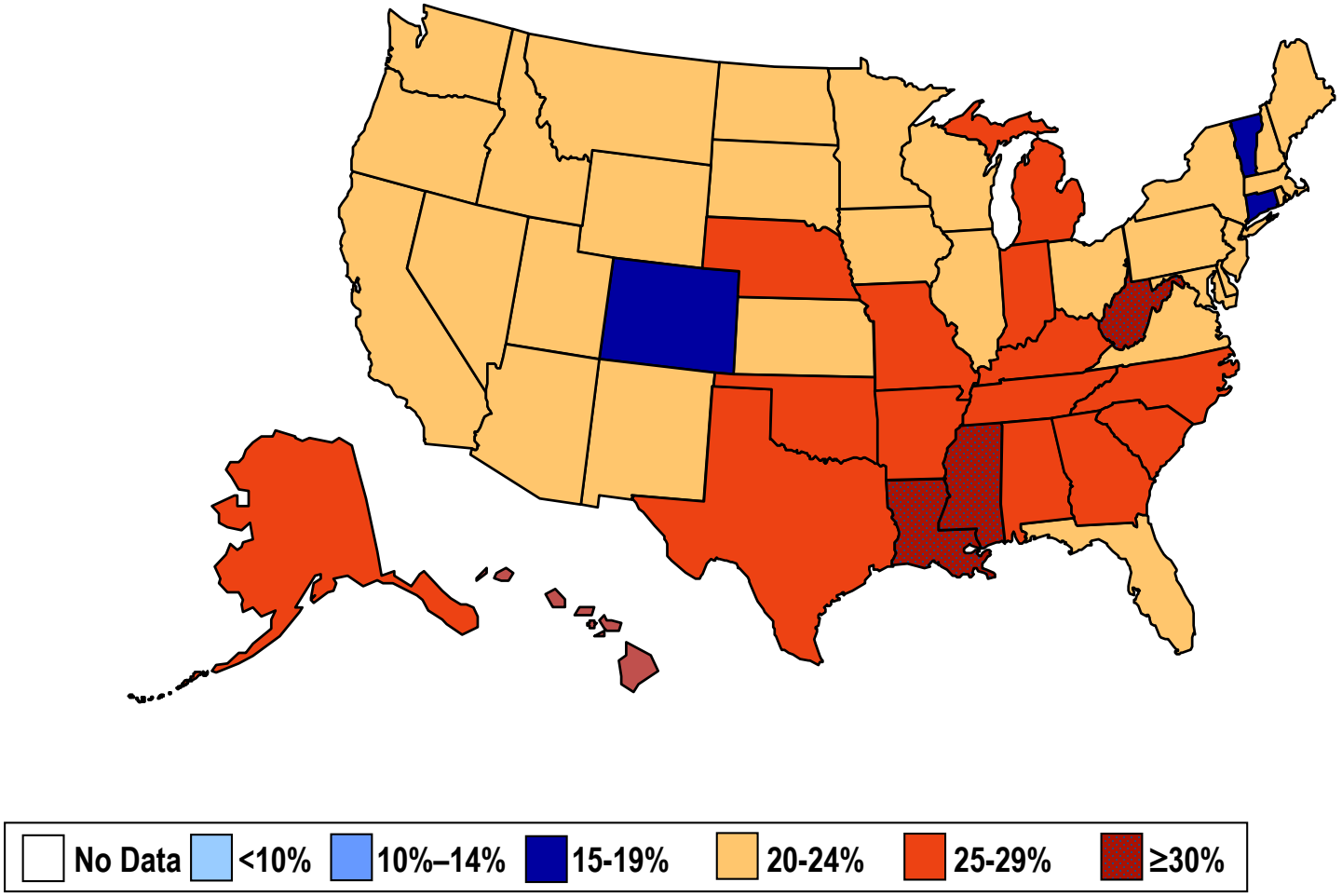
Obesity* in U.S. Adults: 2004

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



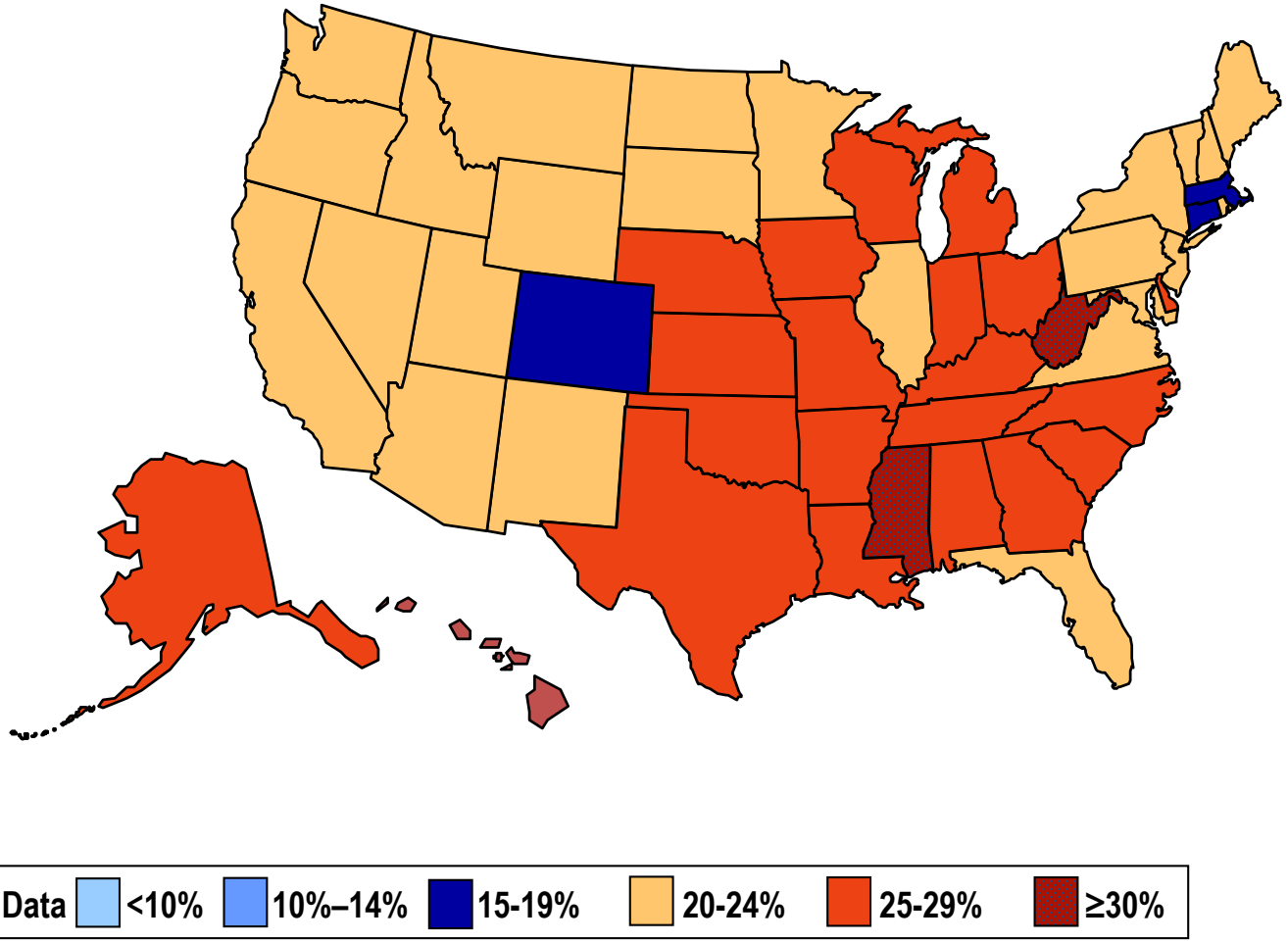
Obesity* in U.S. Adults: 2005

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



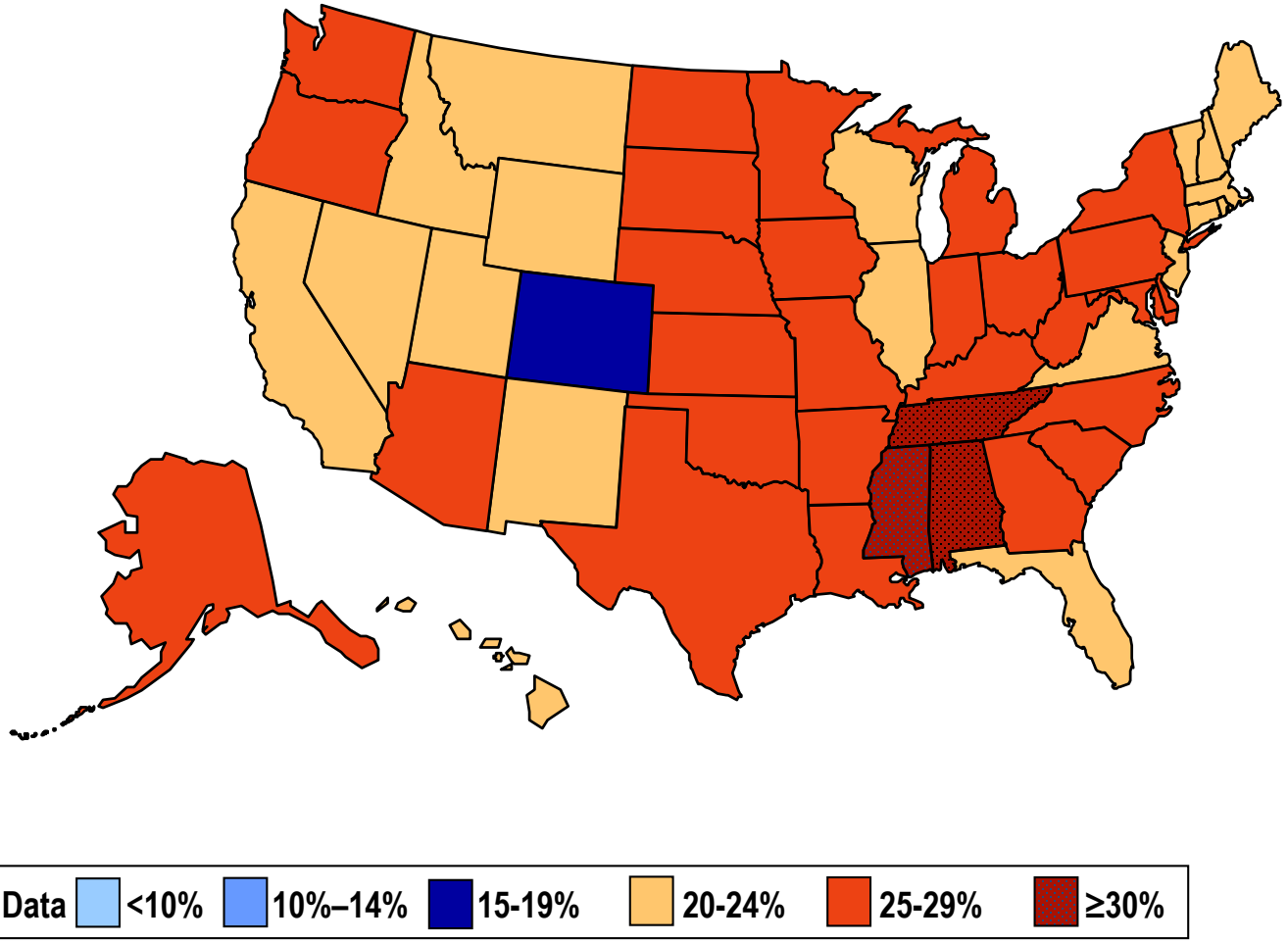
Obesity* in U.S. Adults: 2006

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



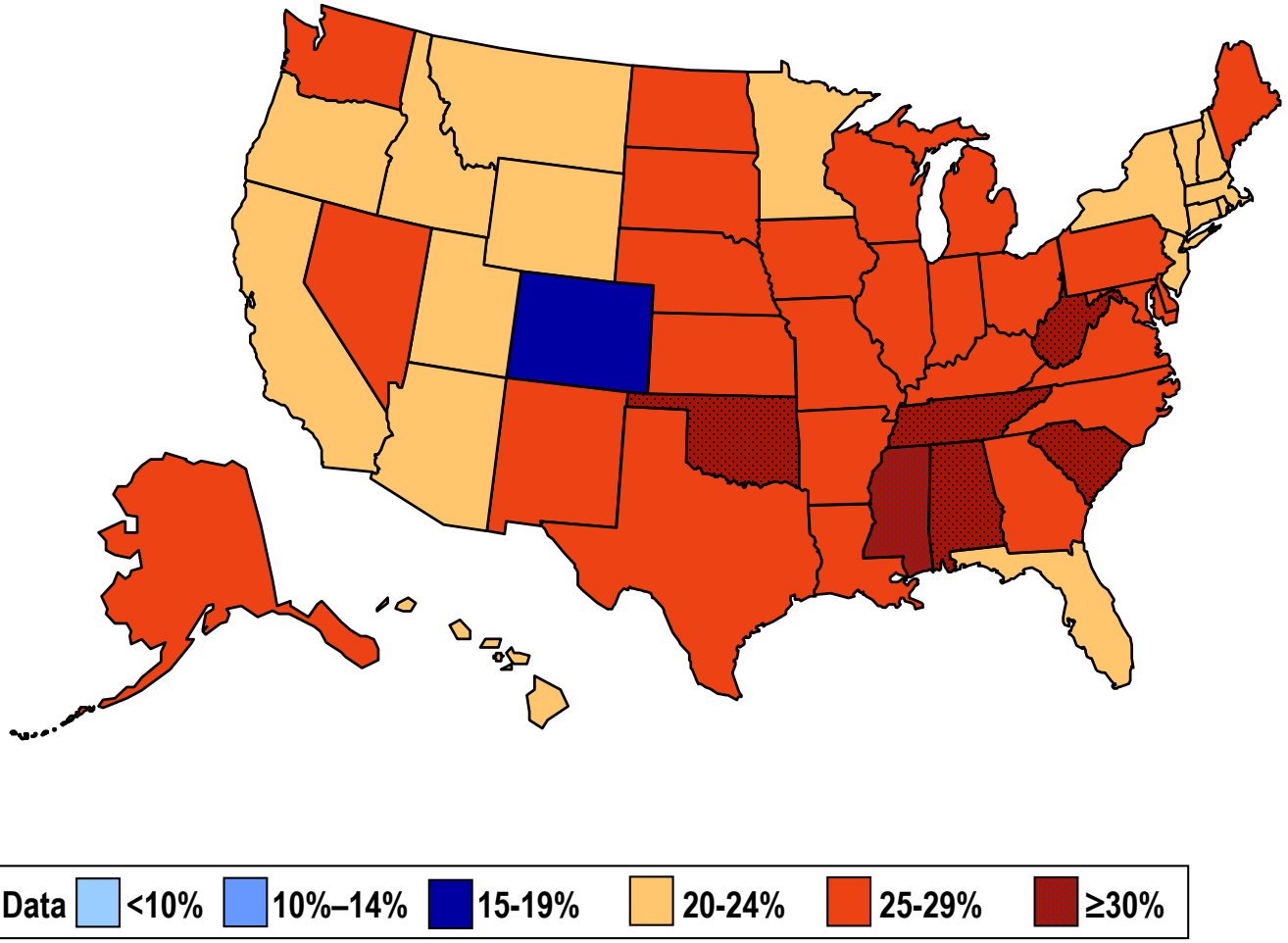
Obesity* in U.S. Adults: 2007

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



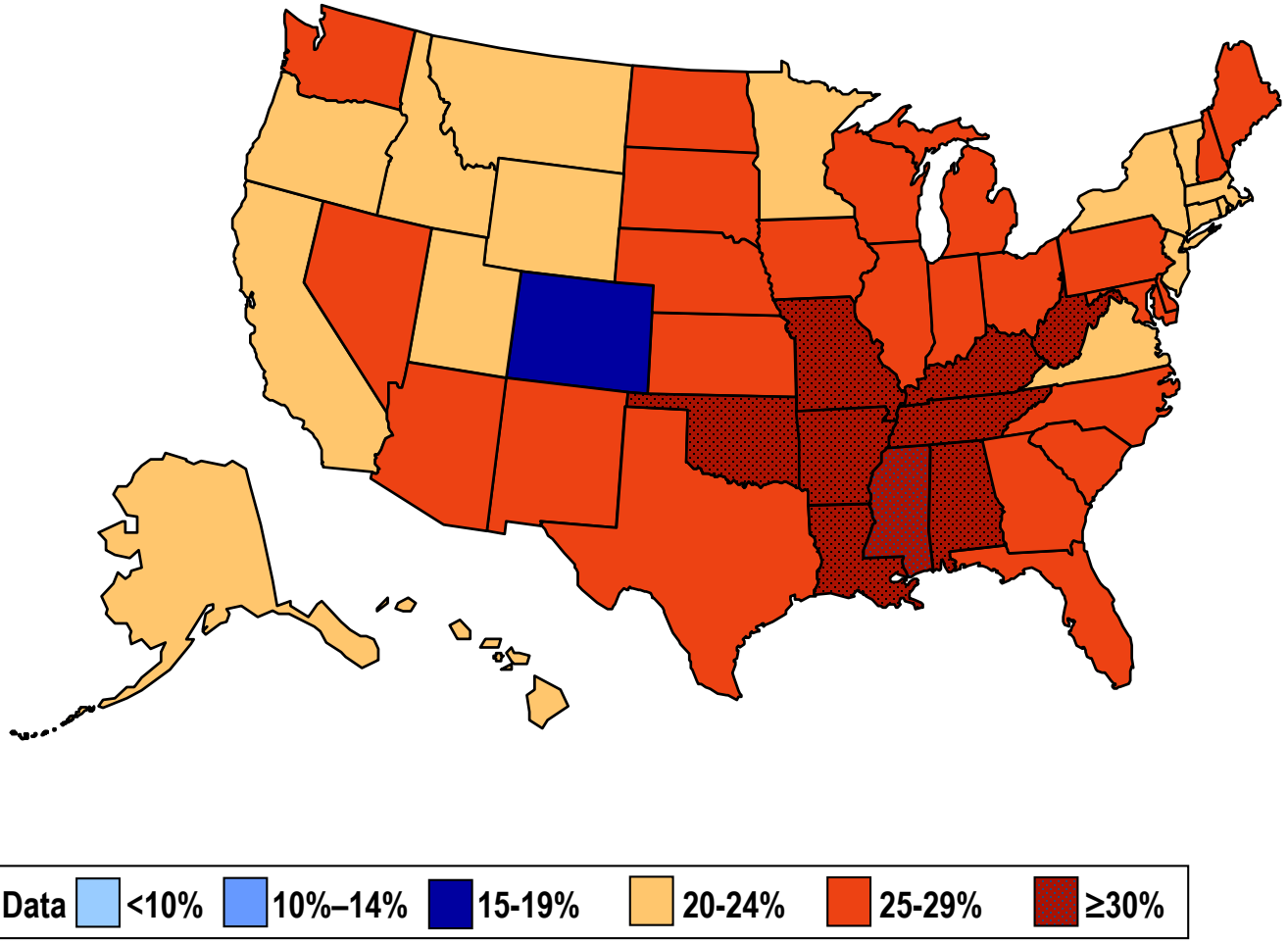
Obesity* in U.S. Adults: 2008

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



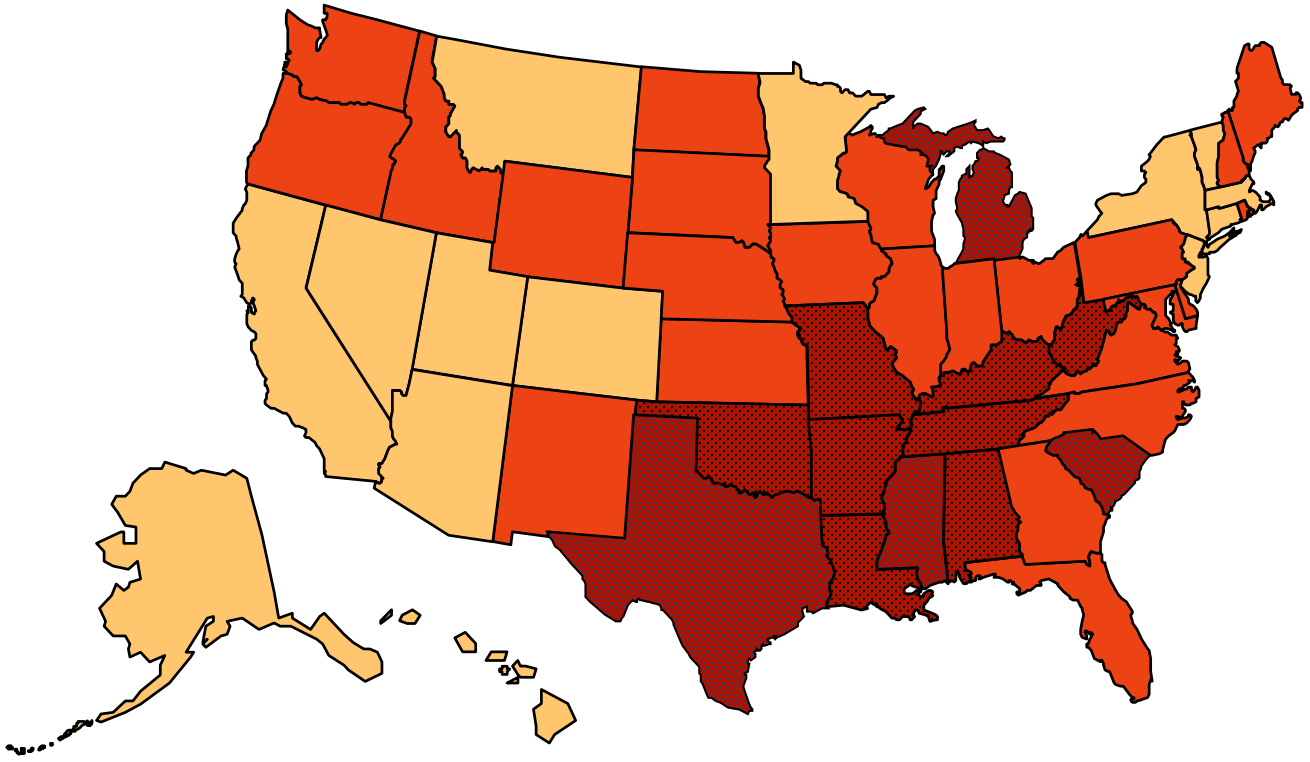
Obesity* in U.S. Adults: 2009

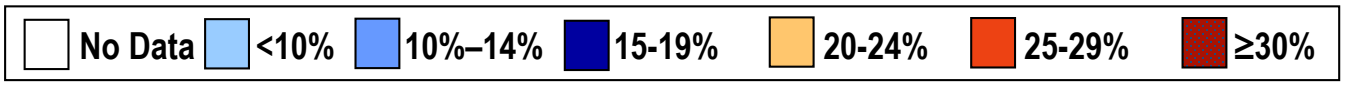
(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)



Obesity* in U.S. Adults: 2010

(*BMI ≥30, or ~ 30 lbs. overweight for 5’ 4” person)

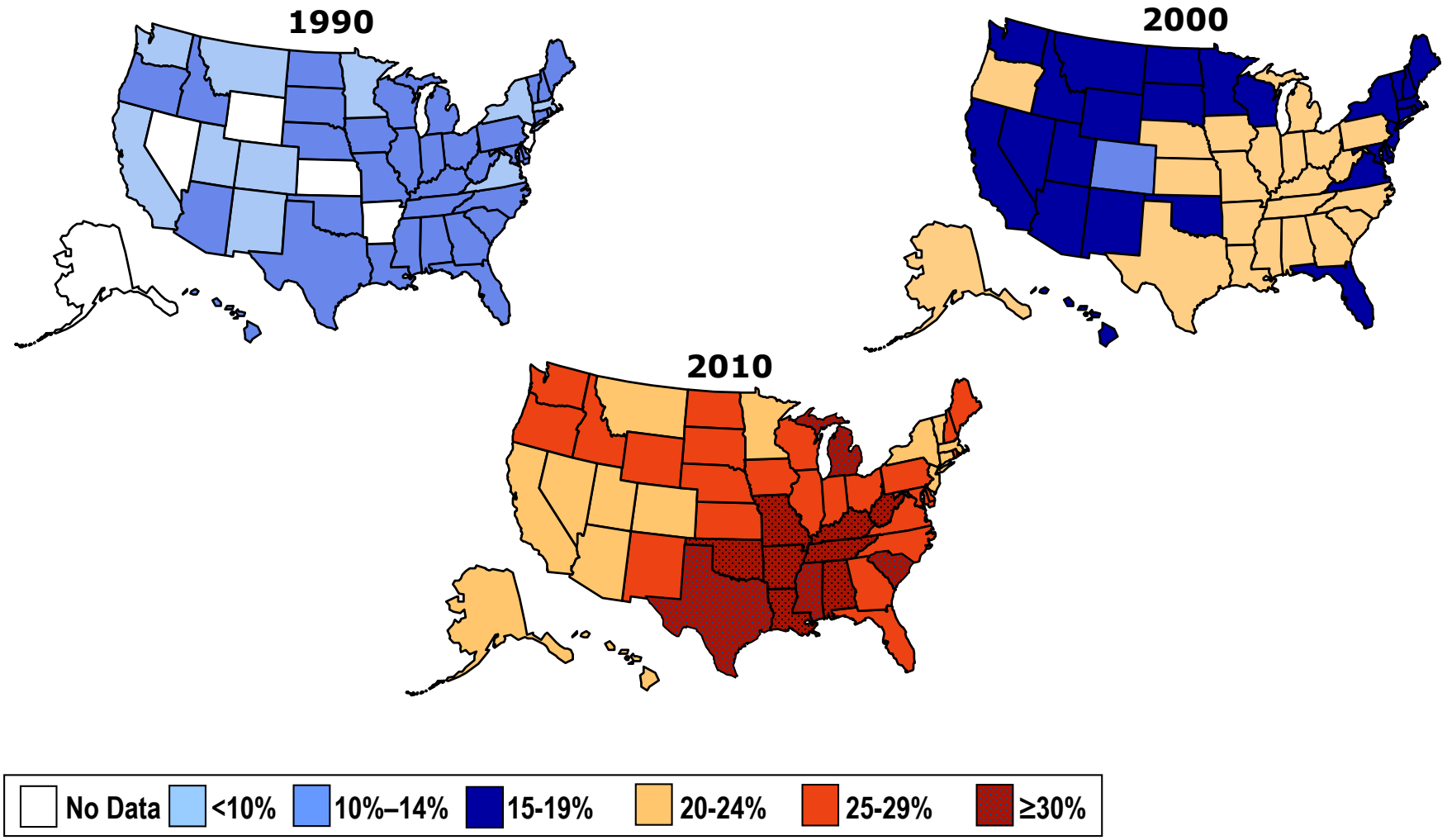




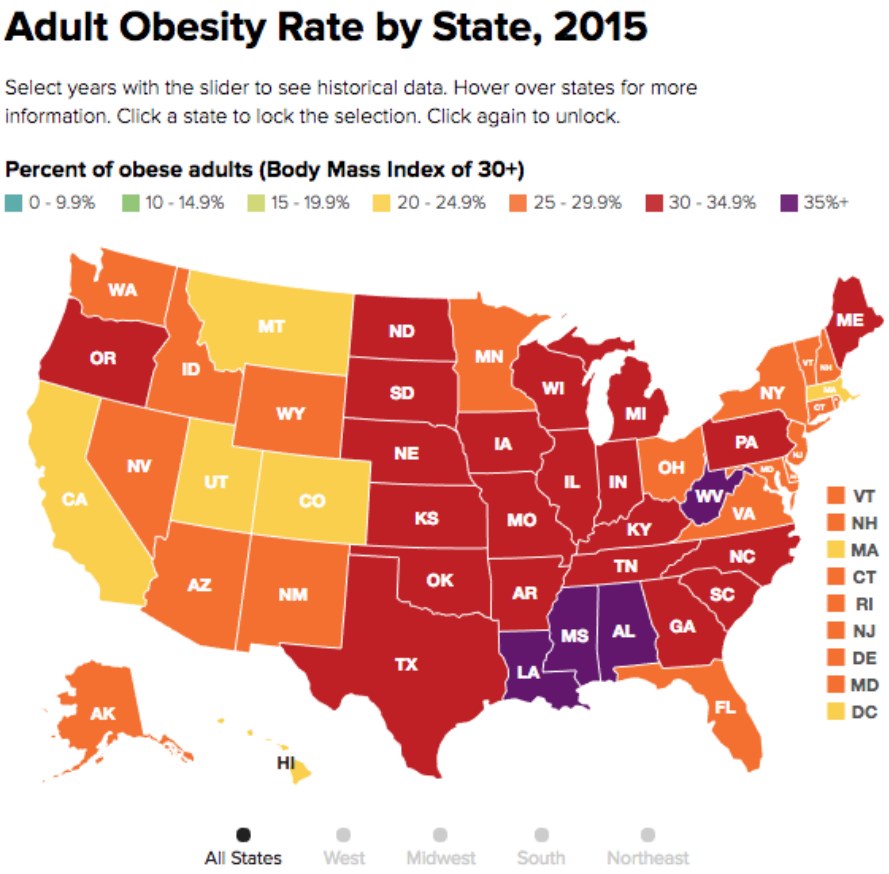
Obesity Trends* Among U.S. Adults

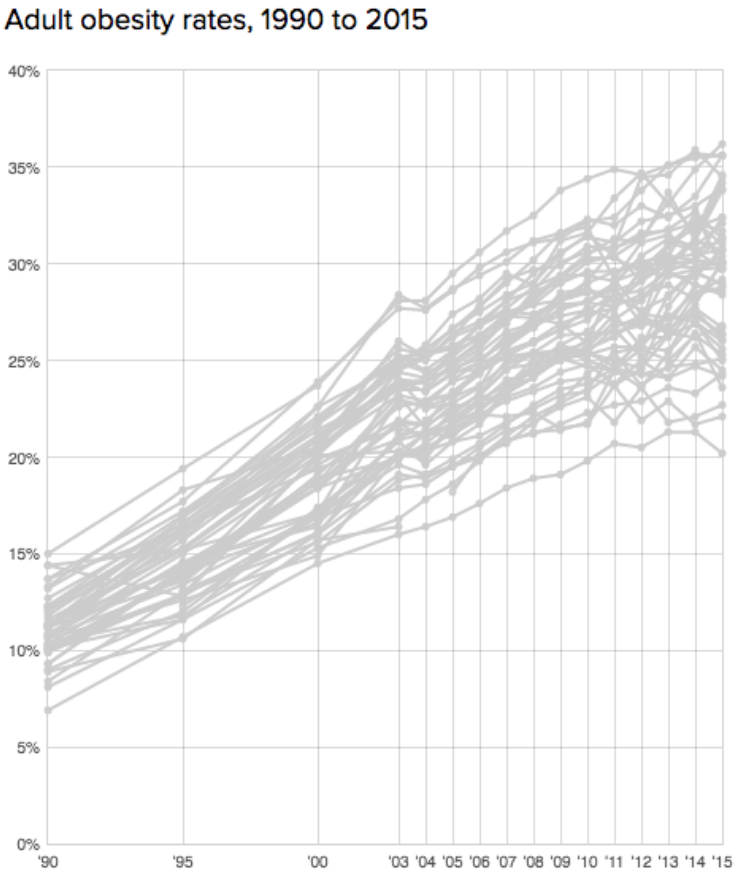
BRFSS, 1990, 2000, 2010

(*BMI ≥30, or about 30 lbs. overweight for 5’ 4” person)



Two Different Visualizations





http://stateofobesity.org/adult-obesity/

Data Visualization

- Visual!
- Intuitive!
- Engages right brain and creativity!
- Emphasizes multivariable thinking!
- Exposes a new world of possibilities!
- Engaging for students of all math abilities!
- Fun!! (for students and teachers!)
- Lots of good examples here: www.personal.psu.edu/klm47/visualization.htm

Intro Stats

Why is there a push now to change?

Technology!!!

We now have the technological ability to collect large datasets and to visualize them.

The Future of Intro Stat: More Accessible

Fewer Prerequisites

NONE of this requires algebra!!!

=> accessible to more students, and more students can succeed!

Also not required:

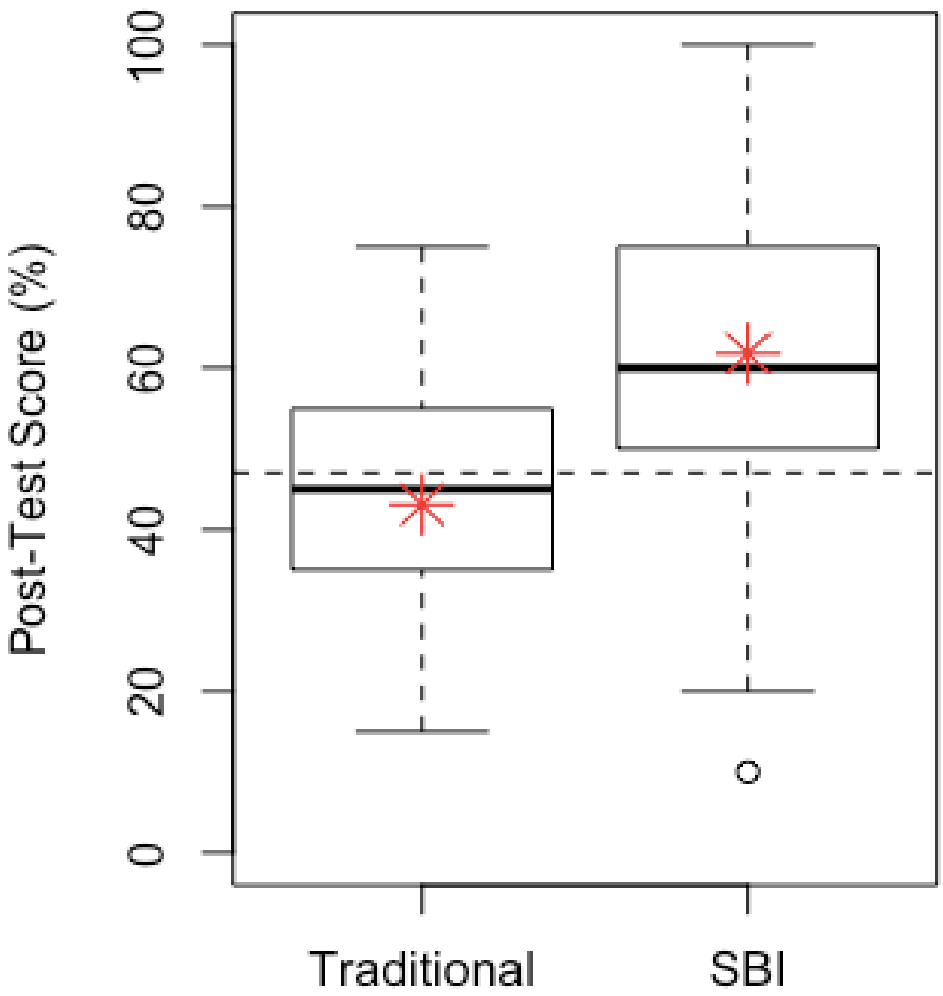
- Formal probability
- Theoretical distributions
- Test statistic or standard error formulas

The Future of Intro Stat: More Effective

More Effective!

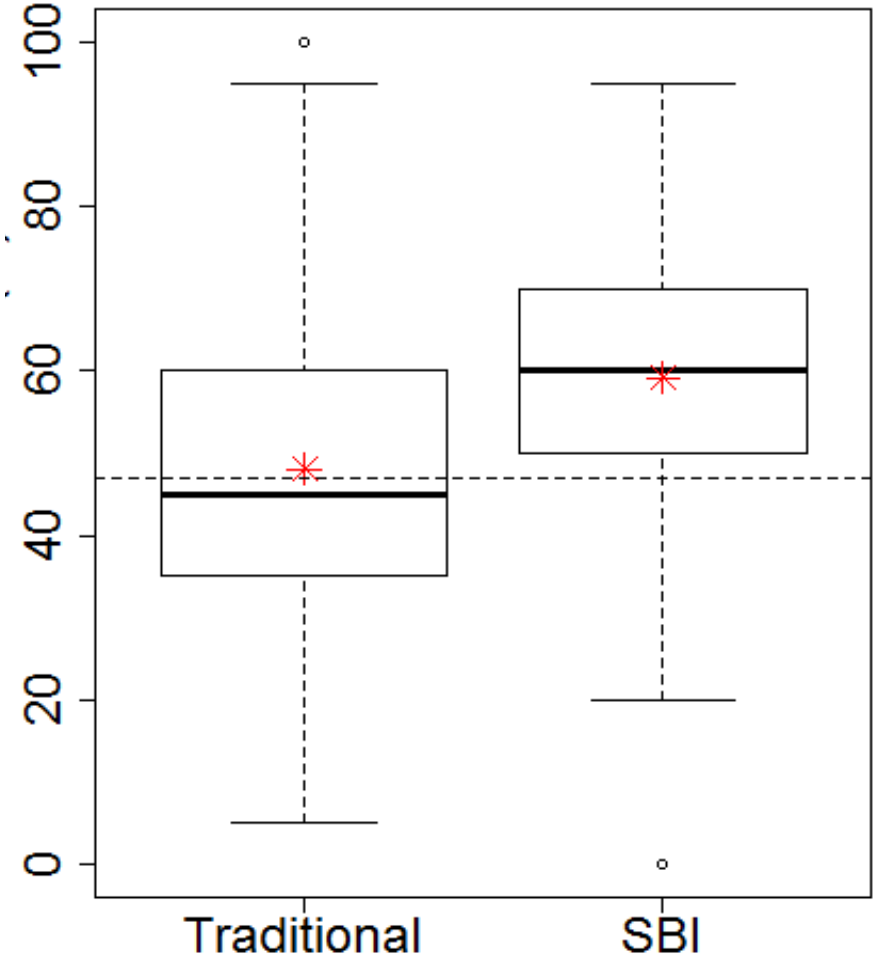
Intro Biostats

$\bar{x}_{SBI} - \bar{x}_{trad} = 18.7$



Intro Stats

$\bar{x}_{SBI} - \bar{x}_{trad} = 11.2$

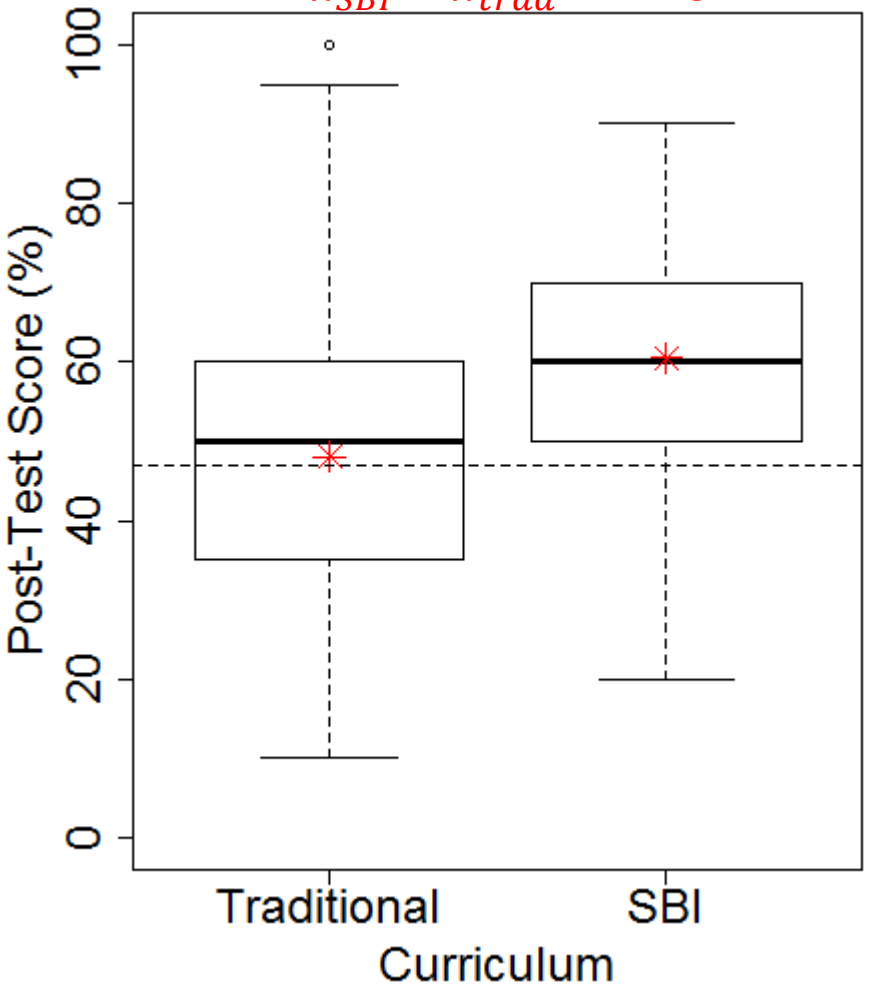


p-value < 0.00001

More Effective!

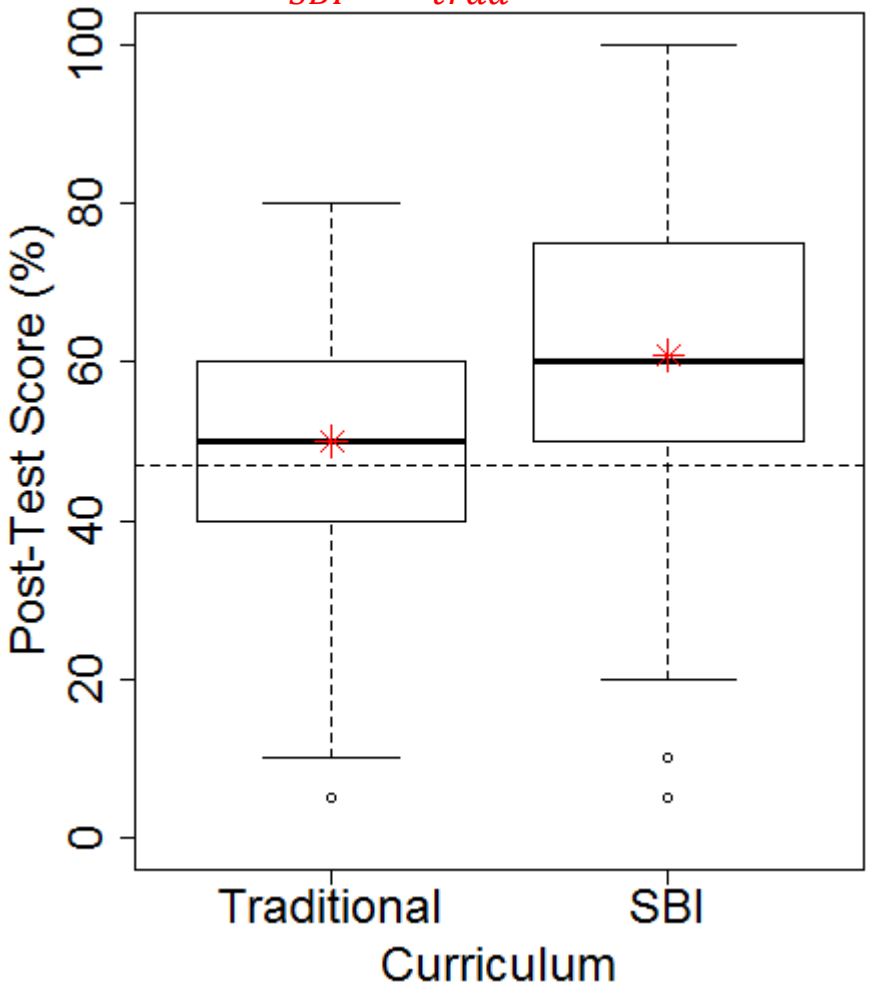
Same instructor A, S17/F17

$\bar{x}_{SBI} - \bar{x}_{trad} = 12.5$



Same instructor B, S17/F17

$\bar{x}_{SBI} - \bar{x}_{trad} = 10.8$



p-value < 0.00001

The Future of Intro Stat:
More Relevant

REAL DATA!!!

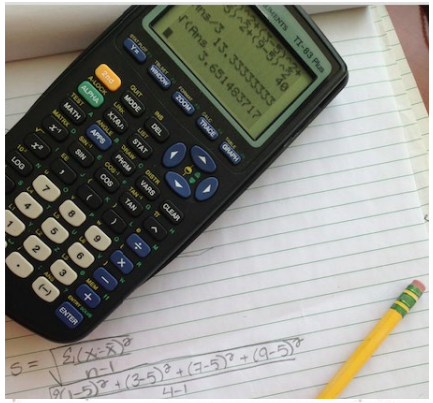
Intro Stats for Who?

- **Students who are good at math**
  - This is one of only three courses the 2015 MAA Curriculum Guide recommends for all math majors
- **Students who are not good at math**
  - A randomized experiment (!) found students needing remedial math had greater success in intro stats than college algebra
- **Students going into statistical or data science**
  - These skills are emphasized in national curriculum guidelines for undergraduate programs in Statistics and Data Science
- **Students who plan to become math teachers**
  - Simulation-based inference recommended by SET 2015
- **Students in fields that analyze data**
  - User disciplines love the emphasis on real data
- **Students taking stats for general education**
  - "statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write." - H.G. Wells

EVERYONE!!!

The Future of Intro Stat: More Fun!

Consider two events A and B, and assume that P(A) = 0.6 and P(B) = 0.5 and P(A∩B) = 0.2.
Find P(A∪B).



$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

TABLE B: *t*-DISTRIBUTION CRITICAL VALUES

	Tail probability *p*											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level *C*											

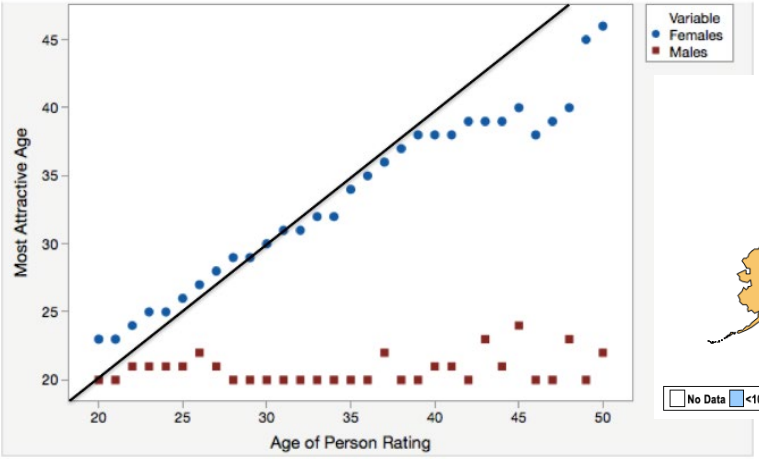
Question of the Day

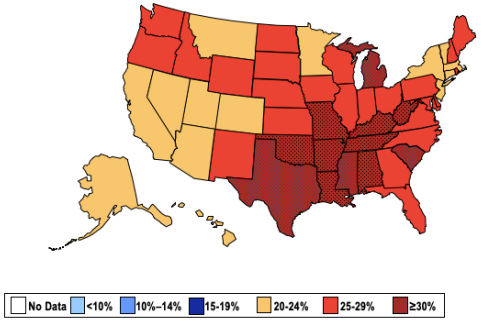
Is cat ownership related to Schizophrenia?

https://www.youtube.com/watch?v=ahd9Cr0U8XU

Statistics: Unlocking the Power of Data — Lock5

We have strong evidence that drinking beer does attract mosquitoes!





Ways to Change

1. Make everything about real **data**
2. Reduce tedious "by-hand" work and let **technology** do the heavy-lifting
3. Focus on interpretation and **concepts**
4. Improve conceptual understanding and reduce reliance on prerequisites with **simulation-based inference**
5. Embrace the power of **data visualization** and multivariable thinking

The Future of Intro Stats

More accessible

More effective

More relevant

More fun!

klm47@psu.edu