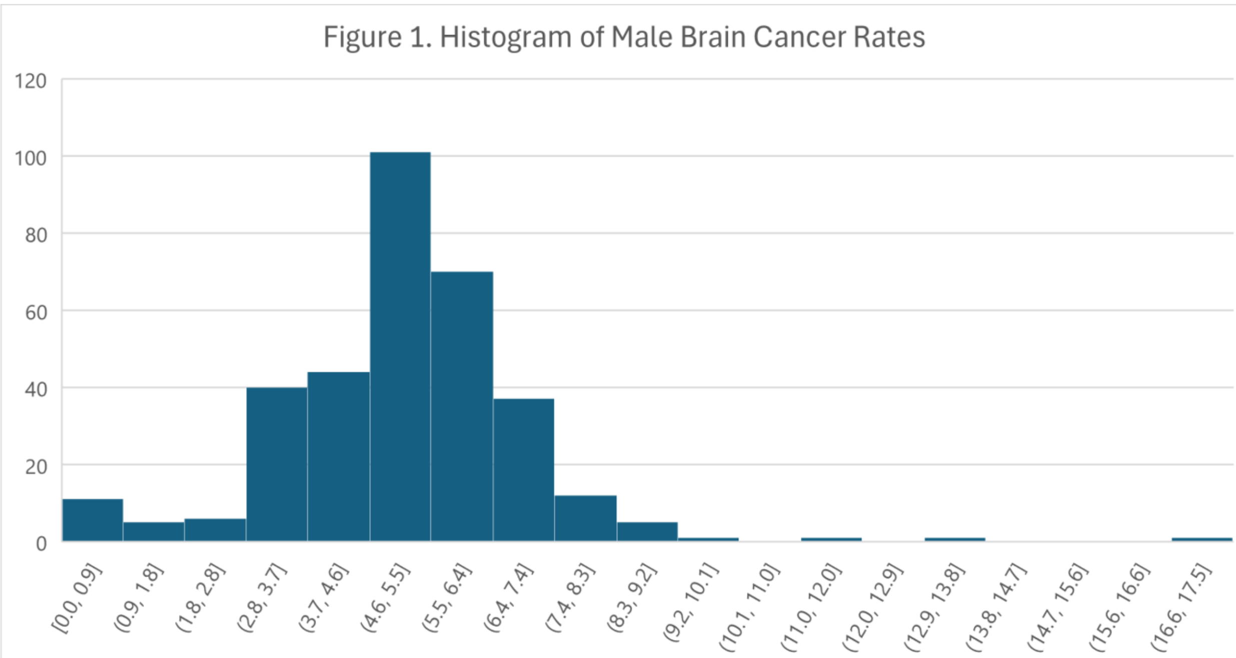
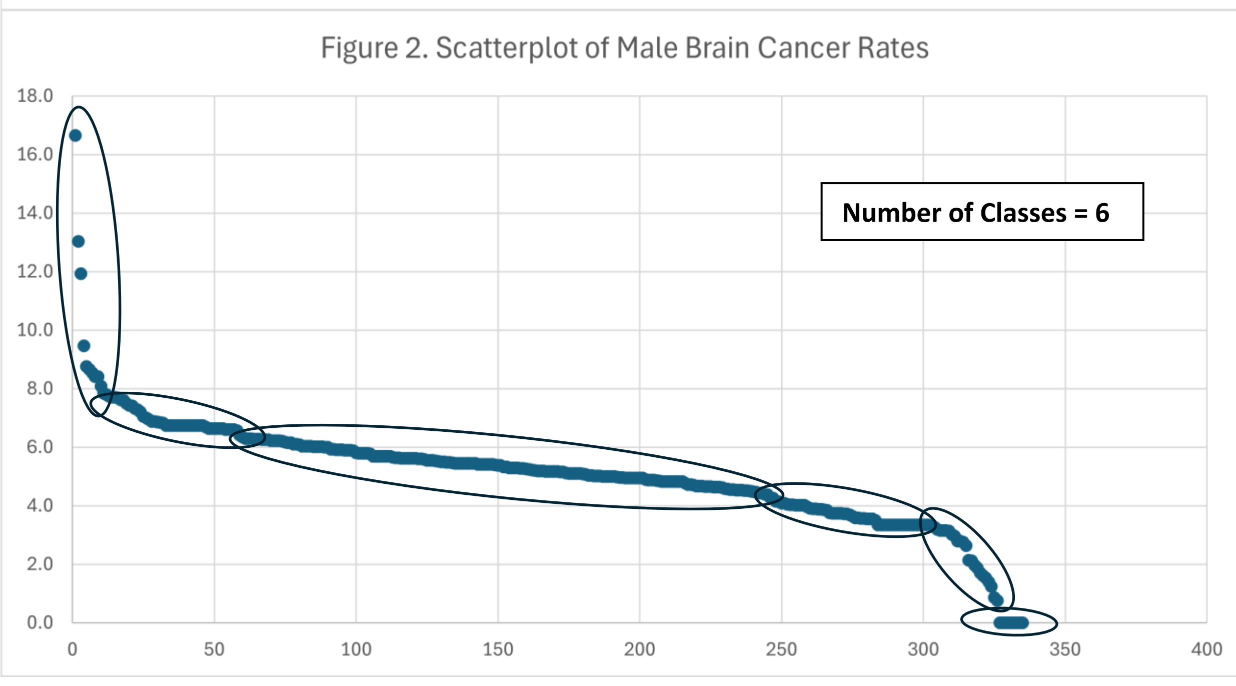
Daniel Ritter
GEOG 6165
February 6, 2024

Week 3 Exercise: Statistical Mapping

1.2 Distribution of Chosen Variable

I chose to use the male brain cancer rate (*BRAM_RATE*) as my variable. *Figure 1* and *Figure 2* depict the distribution of values as a histogram and a scatterplot, respectively.





I divided the male brain cancer rates into six classes based on the scatterplot. The first, fifth, and sixth classes represent outlier datapoints and are clearly separate from the second, third, and fourth classes. While the breaks between the middle classes are less clear, the third class has a slightly shallower slope than the second and fourth classes. I considered combining the fifth and sixth classes into a single class for simplicity's sake, but I decided that the absence of cancer should be separate from presence of cancer, even if it is very low.

1.3 Using CLASSIT to Analyze Data Distribution

OPTIMAL CLASSES AS MEASURED BY ABSOLUTE DEVIATIONS FROM THE MEDIAN

A 10 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**59.970**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
10	3	16.65	11.93	13.03	4.72
9	20	9.47	7.28	7.725	8.42
8	36	7.2	6.41	6.74	3.51
7	46	6.34	5.78	6.03	6.69
6	68	5.68	5.15	5.44	9.5
5	70	5.12	4.43	4.83	12.08
4	31	4.39	3.68	3.97	4.91
3	41	3.64	2.63	3.35	6.22
2	10	2.14	0.86	1.66	3.16
1	10	0.76	0	0	0.76
A 9 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**70.400**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
9	3	16.65	11.93	13.03	4.72
8	20	9.47	7.28	7.725	8.42
7	36	7.2	6.41	6.74	3.51
6	46	6.34	5.78	6.03	6.69
5	68	5.68	5.15	5.44	9.5
4	70	5.12	4.43	4.83	12.08
3	31	4.39	3.68	3.97	4.91
2	46	3.64	1.7	3.35	13.16
1	15	1.62	0	0	7.41
A 8 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**82.480**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
8	3	16.65	11.93	13.03	4.72
7	20	9.47	7.28	7.725	8.42
6	36	7.2	6.41	6.74	3.51
5	46	6.34	5.78	6.03	6.69
4	68	5.68	5.15	5.44	9.5
3	76	5.12	4.14	4.83	15.47
2	71	4.08	1.7	3.35	26.76
1	15	1.62	0	0	7.41

A 7 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					95.960
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
7	3	16.65	11.93	13.03	4.72
6	21	9.47	7.2	7.71	8.93
5	53	7.05	6.15	6.64	11.17
4	83	6.11	5.26	5.62	17.03
3	89	5.24	4.14	4.87	19.94
2	71	4.08	1.7	3.35	26.76
1	15	1.62	0	0	7.41
A 6 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**109.630**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
6	23	16.65	7.28	7.79	26.76
5	54	7.2	6.15	6.64	11.73
4	83	6.11	5.26	5.62	17.03
3	89	5.24	4.14	4.87	19.94
2	71	4.08	1.7	3.35	26.76
1	15	1.62	0	0	7.41
A 5 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**129.020**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
5	74	16.65	6.2	6.74	56.26
4	86	6.17	5.26	5.62	18.65
3	89	5.24	4.14	4.87	19.94
2	71	4.08	1.7	3.35	26.76
1	15	1.62	0	0	7.41
A 4 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**159.410**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
4	90	16.65	6	6.74	67.12
3	155	5.93	4.37	5.17	54.94
2	74	4.28	1.88	3.525	28.24
1	16	1.7	0	0	9.11
A 3 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**195.050**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
3	90	16.65	6	6.74	67.12
2	156	5.93	4.28	5.17	55.83
1	89	4.25	0	3.35	72.1
A 2 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**276.960**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
2	239	16.65	4.49	5.62	197.47
1	96	4.48	0	3.35	79.49
A 1 CLASS MAP WITH TOTAL ABSOLUTE DEVIATIONS OF					**422.710**
Class	# Obs	Largest	Smallest	Median	Abs. Dev.
1	335	16.65	0	5.17	422.71

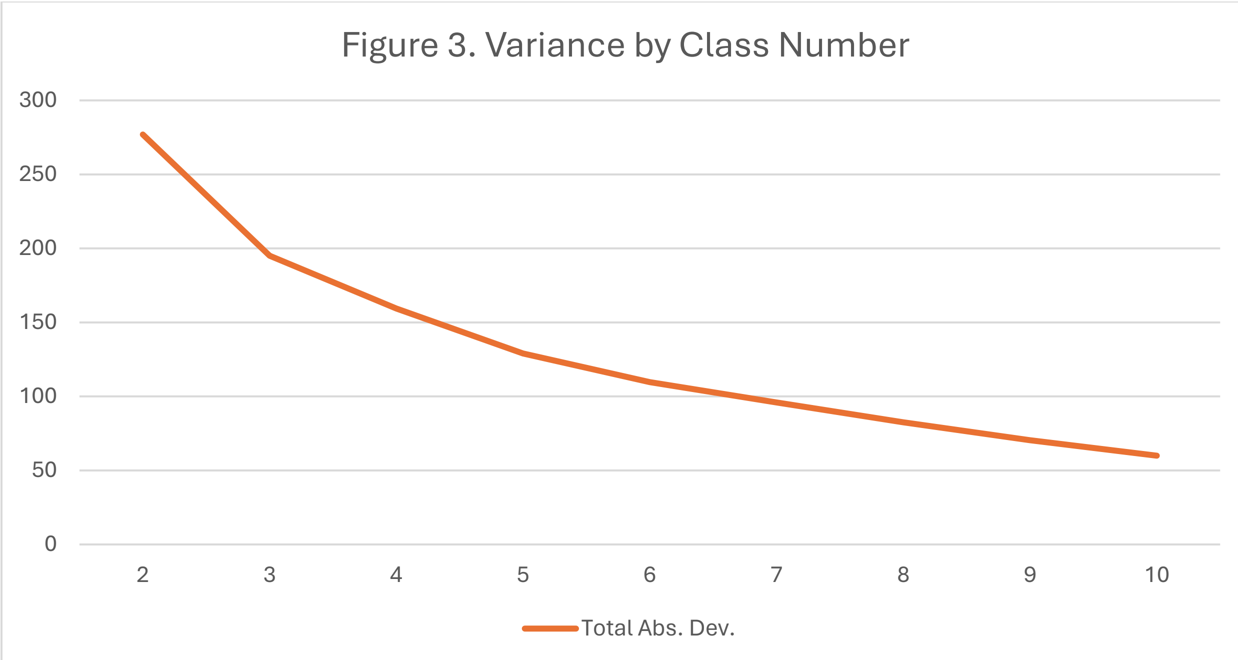
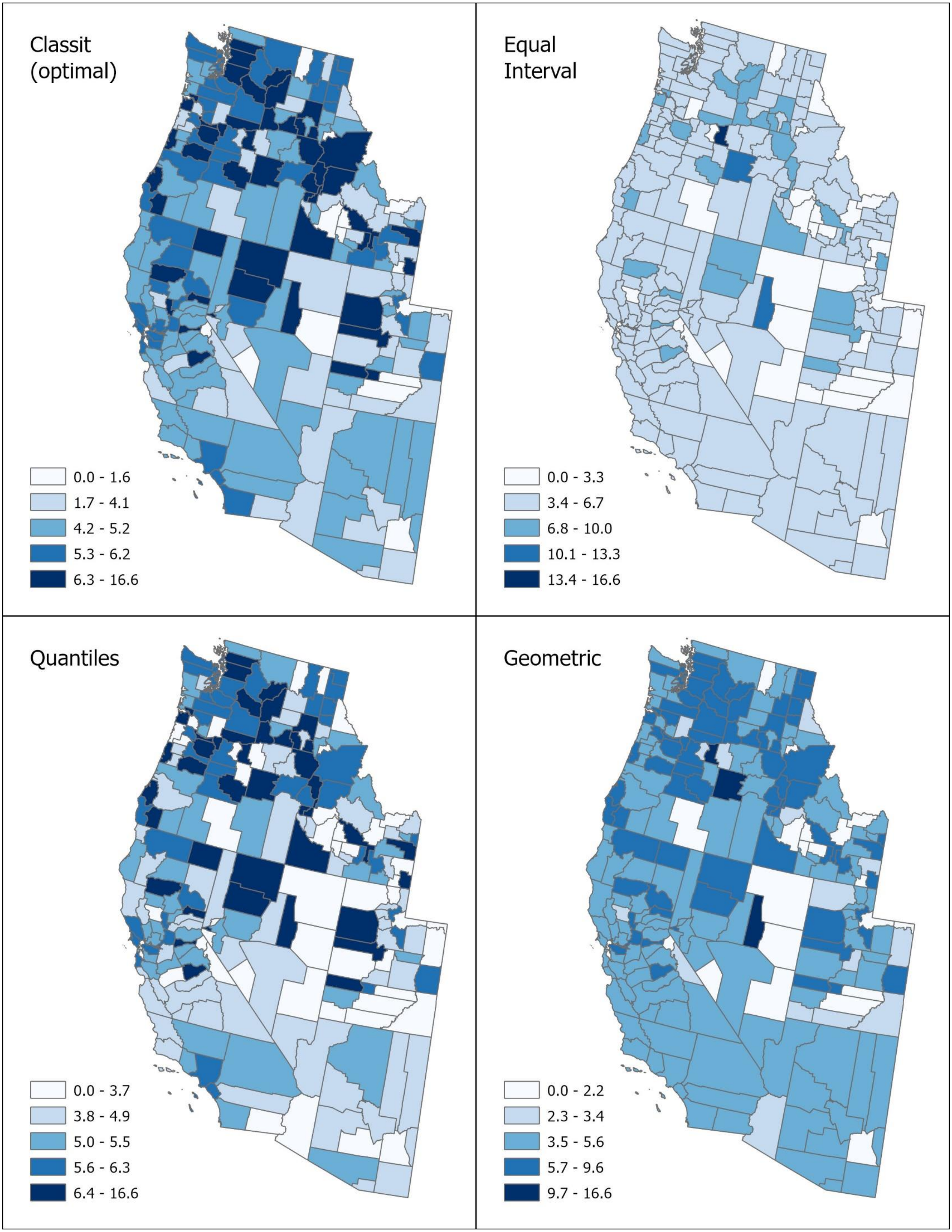


Figure 3 depicts the total absolute deviation, or variance, for each class as defined by CLASSIT. There appear to be two elbows: class 3 and class 5. While the elbow at class 3 is more prominent than the elbow at class 5, I believe that only using three classes would be an oversimplification of the data. Since *Figure 2* indicated there could be five or six classes and *Figure 3* indicated there could be three or five classes, I will use five classes for my maps.

2. Choropleth Mapping

The map below depicts brain cancer rates among males in the western United States using four different classification schemes.



Classit (optimal)
0.0 - 1.6
1.7 - 4.1
4.2 - 5.2
5.3 - 6.2
6.3 - 16.6
Equal Interval
0.0 - 3.3
3.4 - 6.7
6.8 - 10.0
10.1 - 13.3
13.4 - 16.6
Quantiles
0.0 - 3.7
3.8 - 4.9
5.0 - 5.5
5.6 - 6.3
6.4 - 16.6
Geometric
0.0 - 2.2
2.3 - 3.4
3.5 - 5.6
5.7 - 9.6
9.7 - 16.6

4. Map Analysis

Each classification has its own advantages and disadvantages. Both Classit and Quantiles group the counties into bins with relatively similar numbers of observations, which allows for more balanced colors. However, these classifications group the very high cancer rates (e.g., 10.0 and above) with cancer rates that are high but closer to the norm (e.g., 6.0), which deemphasizes the counties with the most extreme brain cancer rates. Both Equal Interval and Geometric identify these counties more clearly, although the wide range of cancer rates means that the fourth and fifth classes for Equal Interval have only three total observations. Further, there appear to be an excessive number of observations in the third class for Geometric. Ultimately, the best classification is the one that fits the purpose of the analysis. Geometric is better for identifying counties with the highest rates, while Classit is better for identifying variation among nearby counties. Brain cancer rates among males appear to be higher in the Pacific Northwest and lower in the Southwest.