Proving *e* is a Chicken Nugget: A Mathematical Look at the Free/Reduced Lunch Program in Colorado

Angela Morrison
Math 6384: Spatial Data Analysis
December 7, 2021

Abstract: The free/reduced lunch program has been helping to ensure school children 18 years of age and younger are provided with nutritional meals at school since 1946. This program has roughly 30 million participants nationwide. Since the pandemic began, the program has had to make a few adjustments to still provide students with these essential meals. In this report, I perform a spatial data analysis of the number of students eligible for this program within the school districts of Colorado to determine if there is any evidence of clustering in the data. This analysis can help with policy recommendations to provide schools with more ways to make sure these meals are reaching students that may still need to be learning remotely or even more COVID safe measures to allow students to eat these meals while attending school in person.

Introduction

It is without question that schools have been greatly affected by the COVID-19 pandemic. From having classes moved to a remote setting, cancelling traditional school events, such as school dances and sports, to administrations having to equip their staff with new tools for these unprecedented times, a lot has changed to accommodate pandemic guidelines. Another aspect of schools that could not escape the clutches of the pandemic was the free/reduced lunch program. This program was established in 1946 and was created to ensure students 18 years and younger were given access to nutritional meals at school [6]. These programs have expanded as the decades have passed such as breakfast programs and even summer lunch programs. These programs which were once a staple in many students' daily lives have had to adjust due to the pandemic.

Due to the free/reduced lunch program being federally funded, there are usually certain restrictions that schools must adhere to such as the type of food that is served with each meal, how the meals are offered, and sometimes even when the meals can be given. With schools having to close, and students being confined to their homes for the most part, many of these restrictions could not be met. The government did relax some restrictions such as allowing schools to bus out certain meals or forcing them to keep strict track of what meals were served to who even if a student was not on the usual list for free/reduced lunches. However, there are still some lasting effects. For example, some families were without an income for a part of the pandemic or are possibly still without a stable income. Without a steady income, some students were unable to pay for even their reduced lunches, and some students that might not have qualified for the program previously were having to rack up unpaid lunch debt while getting these meals throughout the year.

With this in mind, one could ask, “Where are the communities that could have been the most affected by this pandemic in regards to the free/reduced lunch program?” This report aims to give some insight into answering this question. In this project, I will perform some spatial data analysis to identify if there are any clusters of school districts with large amounts of students eligible for the free/reduced lunch program using various tests. In identifying these areas, it should shed light on the districts in greatest need of extra support through policy recommendations which will be shared at the end of this report.

Data

The data for this project comes from the KIDS COUNT data center [8]. This data set includes the number of students eligible for free lunch, the number eligible for reduced lunch, and the number of students eligible for both free and reduced lunch. It also contains information on the percentage of students in each district for these three categories. There is information from 2008 to 2020, so I used the 2020 data since this would be the closest I could get to the start of the pandemic. Note, that in the data itself there are school district names which are removed from this analysis due to the fact that districts have changed throughout time. Meaning, these districts most likely were absorbed by another one during the years available. I would also like to note that there is one district that needed to be removed from the analysis due to a naming issue between the file used for plotting the data and the actual dataset itself. The number of students in each district based on the numbers and percentages given in the original dataset itself was added to the dataset as well.

Figure 1 gives a plot of the school districts in this analysis and is colored based on the percentage of students in that district eligible for the free/reduced lunch program. Figure 2 contains the same information except that it is colored based on the raw number of students in each district eligible for the free/reduced lunch program. The purpose of these two plots is to show that looking at one set of that data does not truly paint the entire picture of what is being investigated. Figure 1 gives a good idea of which schools are having to support a large chunk of their population via the free/reduced lunch program while the second plot merely shows where the highest concentration of students in the program are in the state of Colorado.

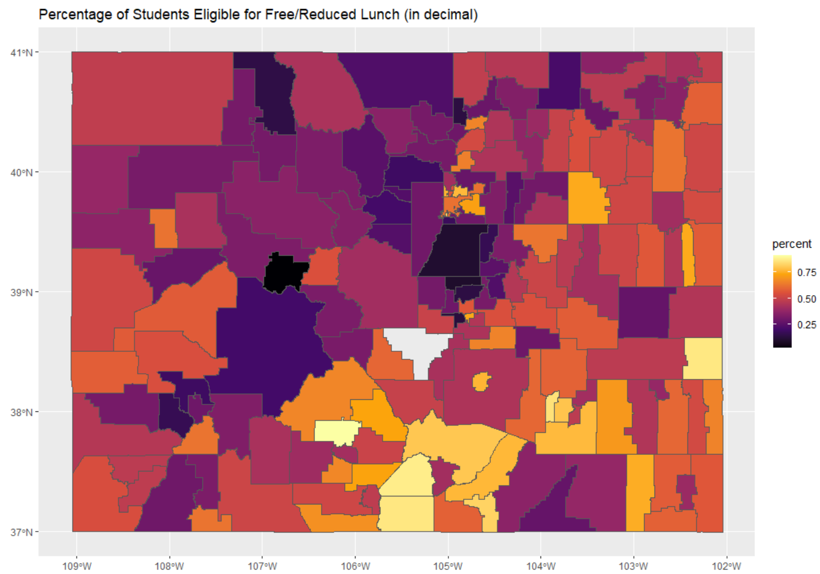

Figure 1: Choropleth plot of percentage of students in districts that are eligible for the free/reduced lunch program.

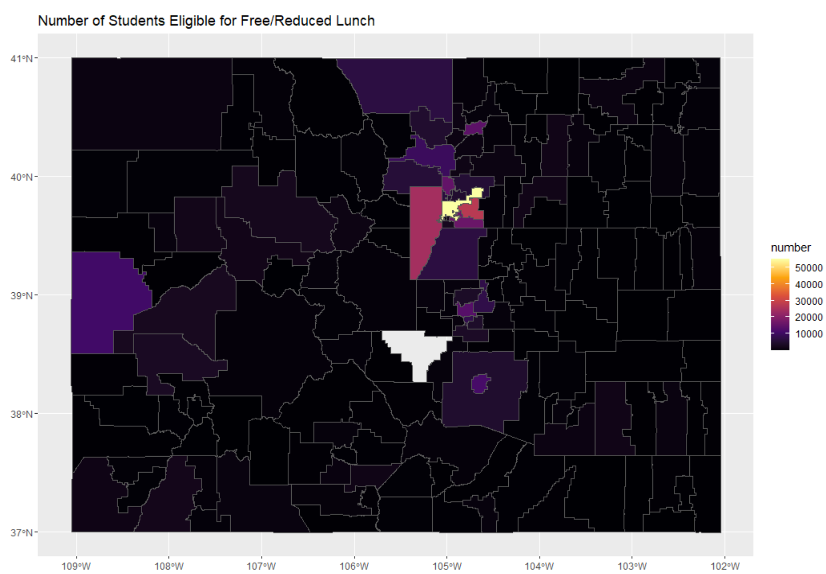

Figure 2: Choropleth plot of raw number of students in districts that are eligible for the free/reduced lunch program.

For this analysis, I will be using the number of students in the free/reduced lunch program for each district for clustering purposes. Doing so allows us to understand if there are certain districts which go against our constant risk hypothesis. That is to say, we will be looking to see if there are any school districts receiving the free/reduced lunch program that is not the same as across all of the districts.

Analysis

For this analysis, I perform three tests: CEPP, Besag-Newell, and Spatial Scan. Each of these tests is meant to detect if there are any clusters in the data and if so, where are they. At the end, I will give a summary of the highlights of each test and whether or not the null hypothesis has failed to be accepted for each test.

CEPP

The null hypothesis for CEPP, the thing we are testing for, is that there is no window with n^* number of students that has significantly more students eligible for the free/reduced lunch program than what is expected under the constant risk hypothesis. Therefore, the alternative hypothesis is that there exists a window with n^* number of students that has significantly more students eligible for the free/reduced lunch program than what is expected under the constant risk hypothesis.

For my analysis, I chose to test a varying number of values for n^* in order to get an idea of what is happening with the data under different conditions. I specifically chose n^* to be 5,000, 10,000, 15,000, and 20,000. This was done to capture what might be happening in both smaller school districts as well as some of the larger ones. Due to the vast difference between the smallest and largest school district values, these were reasonable values to pick in order not to have every district nor the entire state considered a possible cluster. Figures 3 through 6 show the results of the clusters found from CEPP using these $n*$ values.

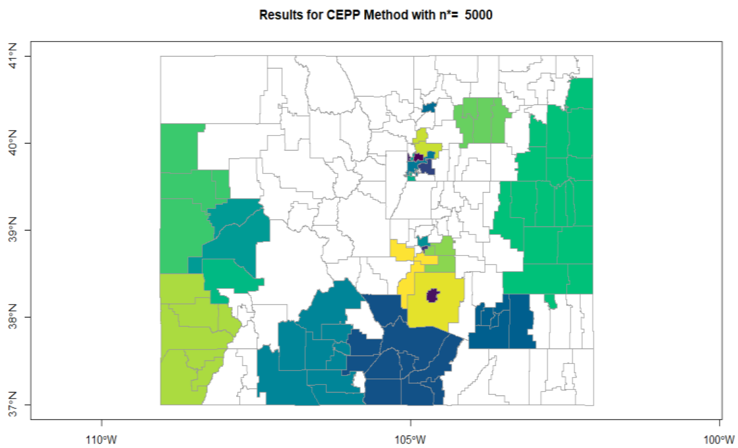

Figure 3: Clusters from the CEPP test when $n* = 5{,}000$

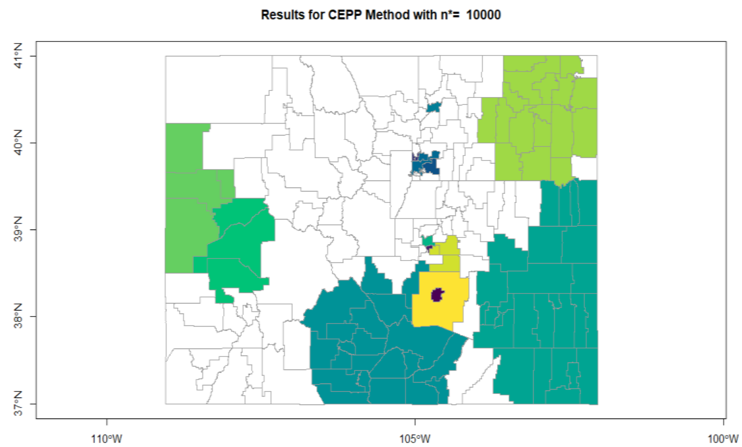

Figure 4: Clusters from the CEPP test when $n* = 10{,}000$

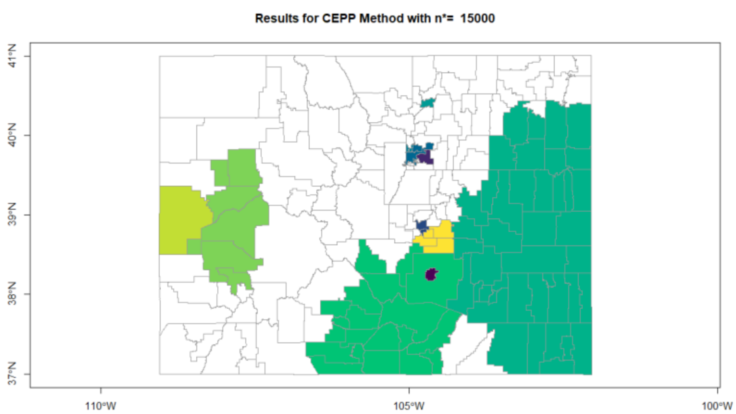

Figure 5: Clusters from the CEPP test when $n* = 15{,}000$

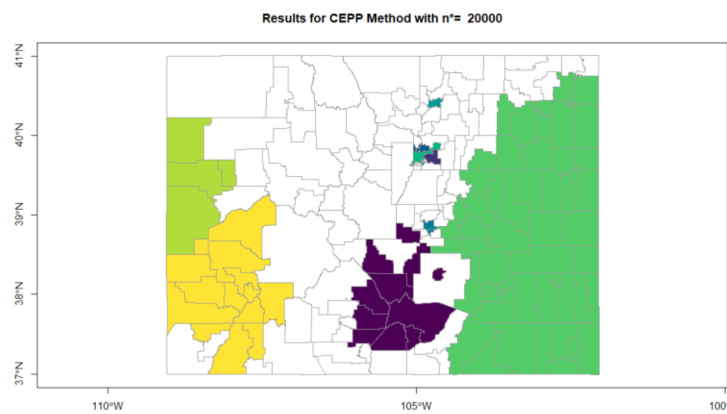

Figure 6: Clusters from the CEPP test when $n* = 20{,}000$

<u>Besag-Newell</u>

The null hypothesis for the Besag-Newell test is there is no window with $c*$ number of students eligible for the free/reduced lunch program that is significantly more compact in terms of total number of students than what is expected under the constant risk hypothesis. The alternative hypothesis would then be that there exists a window with $c*$ number of students eligible for the free/reduced lunch program that is significantly more compact in terms of total number of students than what is expected under the constant risk hypothesis.

As with the CEPP test, I chose a few different values of c^* to test. These values were based on what seemed fitting for the number of students eligible for free/reduced lunch in an average school district since once again the values were so drastic in the dataset itself. The c^* values were 500, 2,500, 5,000, and 10,000 and the plots for these values are in Figures 7 through 10 respectively.

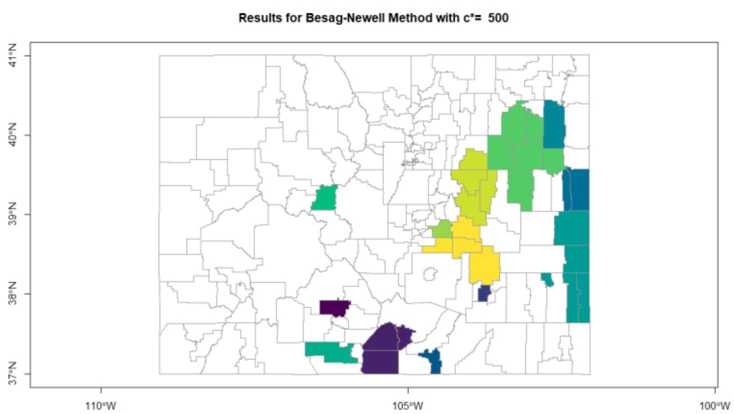

Figure 7: Clusters from the Besag-Newell test when $c^* = 500$

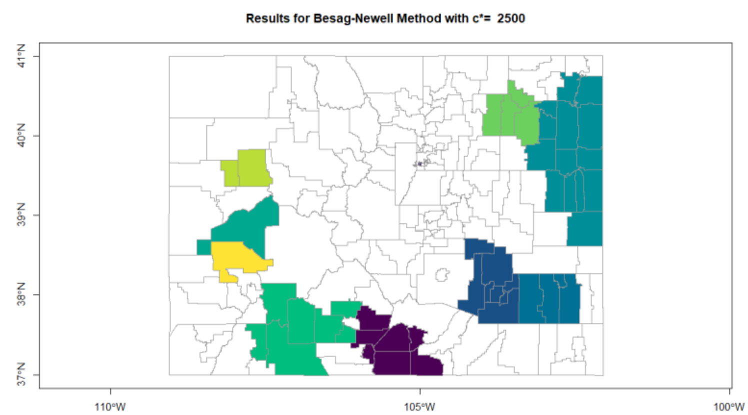

Figure 8: Clusters from the Besag-Newell test when $c^* = 2{,}500$

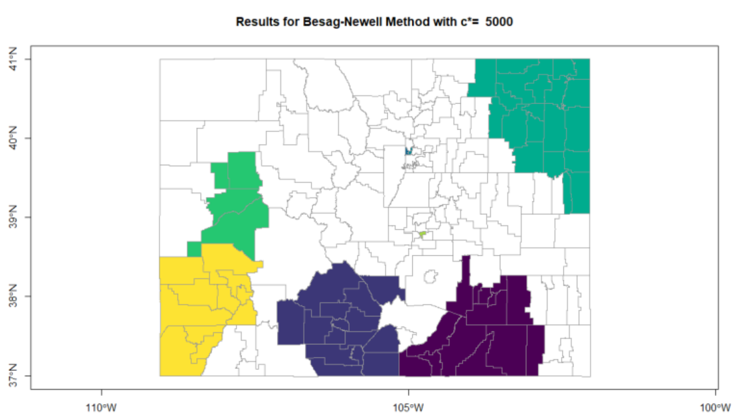

Figure 9: Clusters from the Besag-Newell test when $c^* = 5{,}000$

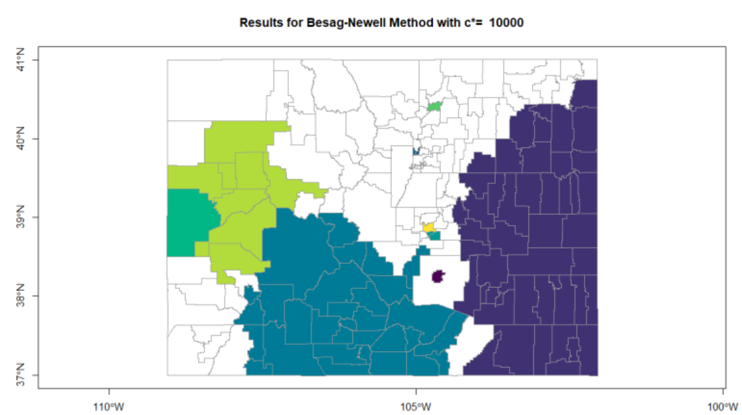

Figure 10: Clusters from the Besag-Newell test when $c^* = 10{,}000$

<u>Spatial Scan</u>

The null hypothesis for the Spatial Scan method is that the most likely cluster of students eligible for the free/reduced lunch program is consistent with what is expected under the constant risk hypothesis. The alternative hypothesis would then be that the most likely cluster of students eligible for the free/reduced lunch program is more extreme than what is expected under the constant risk hypothesis.

While the CEPP and Besag-Newell tests require a certain number of the population or cases to be in each window, the Spatial Scan method just simply has an upper bound on the population of each window. This is usually represented as some sort of percentage of the population. For my analysis, I chose the values for the upper bound to be 0.001, 0.001, 0.01, and 0.1. Again, this was to stop from too many individual districts being considered possible clusters as well as the entire state being a cluster. The plots for these values are Figures 11 through 14 respectively.

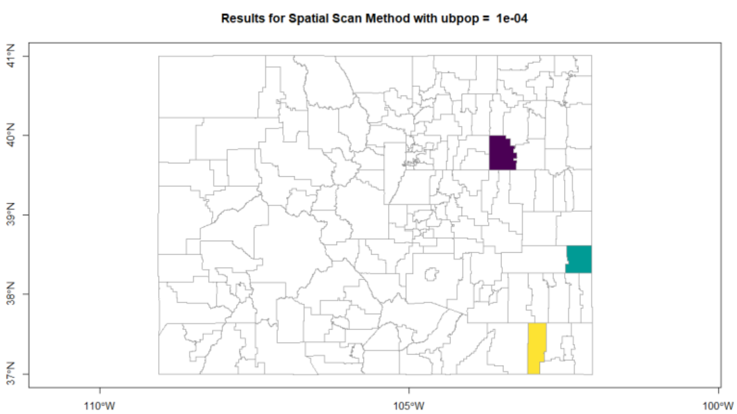

Figure 11: Clusters from the Spatial Scan test with the population bound as 0.0001.

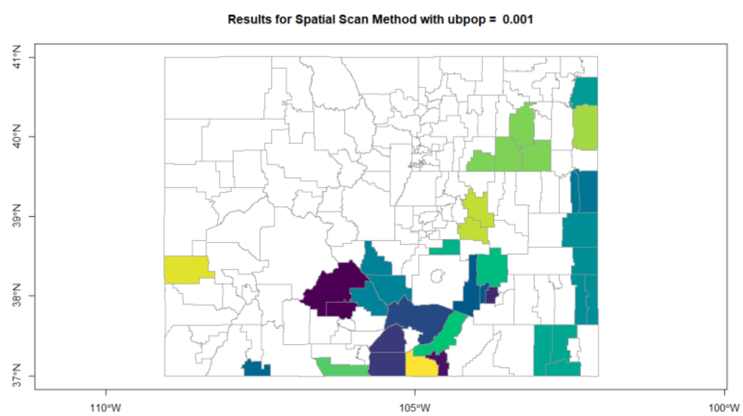

Figure 12: Clusters from the Spatial Scan test with the population bound as 0.001.

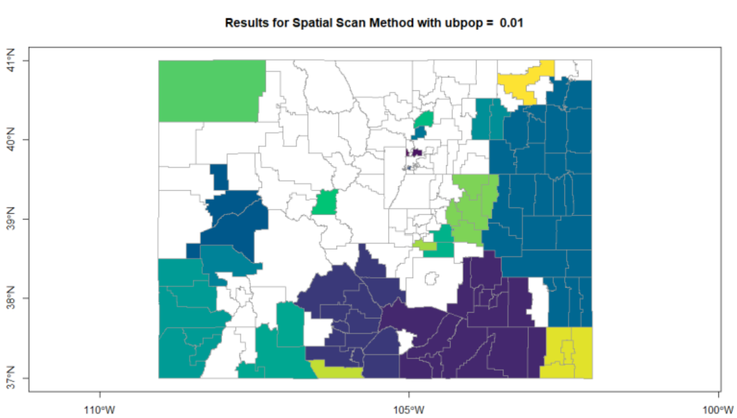

Figure 13: Clusters from the Spatial Scan test with the population bound as 0.01.

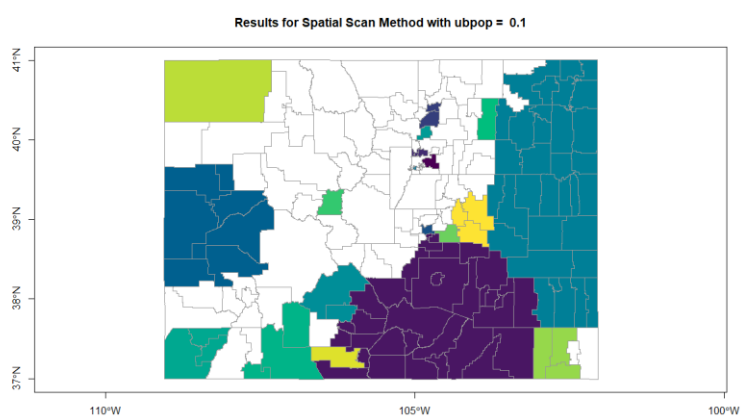

Figure 14: Clusters from the Spatial Scan test with the population bound as 0.1.

Results

CEPP

For the CEPP test, you can find the relevant information summarized in Table 1. From this, it can be seen that for each value of *n** there were statistically significant clusters that were found. This means that we reject the null hypothesis in favor of the alternative hypothesis. In short, there is some window with *n** cases, for each of these *n** values, where there are significantly more students eligible for the free/reduced lunch program than what is expected under the constant risk hypothesis. Table 1 shows that the clusters found for the varying values of *n** are all statistically significant, and one can also see that while the number of clusters overall is decreasing, the sizes of them are increasing. Due to this, one should focus on figures for the smaller values of *n** to get a more precise location of which districts are significant.

n* **Value**	**Number of Clusters**	**Largest Cluster Size**	**Largest p-value**
5,000	23	18	0.002
10,000	14	27	0.002
15,000	11	42	0.002
20,000	9	53	0.002

Figure 6 gives an idea of roughly where something might be occurring within the data. However, with the *n** being such a large value it is not very useful to look at just those colored clusters and say that those are the exact places one should look at. It does show that something is for sure happening in the west part of the state, the middle south part, the eastern part, and around Denver itself. This pattern can still be seen as the value of *n** decreases through Figures 3 through 5. Figure 3 provides some of the best information since the sizes of the clusters are small enough that we get a more targeted location of what districts have more students eligible for the free/reduced lunch program than expected under the constant risk hypothesis. These areas are the ones near the Denver area, the middle southern portion of the state, and northeastern area of the state, and the middle west of the state. Overall, this test is not the only one I considered for my overall conclusions for this report.

Besag-Newell

Similar to CEPP, the relevant information from this test is summarized in Table 2. Due to all of the p-values for the varying *c** values being less than 0.05, the null hypothesis is rejected in favor of the alternative hypothesis. For all values of *c** there exists a window with *c** number of students eligible for the free/reduced lunch program that is significantly more compact in terms of total number of students than what is expected under the constant risk hypothesis. Table 2 shows a pattern of the number of clusters decreasing while the sizes of the clusters increase. Therefore, one can take Figure 10 to give a general idea of where the clusters of these districts are, and as the value of *c** decreases a more precise subset of districts can be seen to be part of these statically significant clusters.

*c** Values	Number of Clusters	Largest Cluster Size	Largest p-value
500	13	6	0.0035
2,500	10	14	0.00099
5,000	8	20	0.000025
10,000	9	48	0.000053

Figure 10 shows similar results to what was seen in the CEPP test, however, as the value of *c** decreases, some of the districts that remain significant in the CEPP test begin to disappear in the Besag-Newell results. In particular, the only regions that I conclude have clusters are near the southern middle part of the state and the northeastern part of the state.

Spatial Scan

Last but not least are the results of the Spatial Scan test. Table 3 summarizes the important results of this test. Since all of the clusters for the varying population bounds are statistically significant, the null hypothesis is once again rejected. Therefore, the conclusion is that for all values of the upper bound on the population, the most likely cluster of students eligible for the free/reduced lunch program is more extreme than what is expected under the constant risk hypothesis.

Again, as the upper bound on the population becomes smaller there are both less clusters and smaller cluster sizes. However, for this test, the figure that is the most telling is Figure 12. Figure 11 I believe to be too small of an upper bound to gain any useful interpretations while Figure 14 includes too much of the state to get an idea of which districts in particular are truly effected or clustered together. Figure 12, in relation to the other figures for the Spatial Scan test, has both a small number of clusters as well as small cluster sizes. Notice that these clusters are all in some of the same spots as with the CEPP and Besag-Newell tests. In particular, the clusters are near the southern middle part of the state, the northeastern part, and there is a single cluster on the west side of the state. These are the clusters I would conclude as being relevant results from the Spatial Scan test.

Population Upper Bound (as a percent)	**Number of Clusters**	**Largest Cluster Size**	**Largest p-value**
0.01%	3	1	0.003
0.1%	21	4	0.003
1%	21	26	0.001
10%	19	34	0.003

Conclusions

Based on the results of the three tests, there is certainly evidence of clusters in the free/reduced lunch program data. In particular, the clusters are located in the southern middle part of the state and the northeastern part of the state. While there are a few other cluster locations that appeared in some of the tests, because they did not appear in all three tests, I do not feel they can be treated as true cluster locations. This is part of the reason for running varying tests on the data.

Policy Recommendations

There are a few policy recommendations that could be given in general for schools struggling with the free/reduced lunch program such as applying for grants or looking into additional programs that the federal government can offer additional funding for [3]. However, based on the locations of the clusters for the districts with a statistically significant number of students that are part of the free/reduced lunch program I have some more specific policy recommendations.

For starters, the cluster locations I mentioned in the Conclusion section are in more rural parts of the state. Typical rural locations have a harder time making their application forms accessible especially during the pandemic. One recommendation could be to require schools to meet certain application requirements [9]. For example, they could make it so that they must offer the forms in multiple languages, or offer a night class that walks people through what the form is asking for, etc. This could be something that the state itself can insist on or they could try to incentivize schools to do it on their own by offering additional funding for doing something similar to this.

The next recommendation I am making is to see if the schools in those districts are interested in community eligibility [9]. This program is something that is offered by the federal government. In short, it allows schools to make lunches free to all students, and schools are then reimbursed based on the number of students in the school that are eligible for free/reduced lunch in the first place. These schools could coordinate together in using available resources to both provide free meals for all students and in conjunction with the community eligibility they could work out a system of making up for the “lost” funds they have if their entire school is not eligible for free/reduced lunch.

Overall, these recommendations won’t necessarily work for all schools in the regions identified by the tests run in this report, but this gives a starting point of where to start looking in terms of schools that require more assistance.

References

1. Editor, N. C. E. S. B. (2015, April 16). NCES blog. IES. Retrieved November 30, 2021, from https://nces.ed.gov/blogs/nces/post/free-or-reduced-price-lunch-a-proxy-for-poverty.
2. Emergency feeding. CDE. (n.d.). Retrieved November 30, 2021, from https://www.cde.state.co.us/nutrition/nutriemergencyfeeding.
3. Grants and awards. CDE. (n.d.). Retrieved November 30, 2021, from https://www.cde.state.co.us/nutrition/nutrigrantsandawards.
4. Home delivery for SFS sponsors | CDE. (n.d.). Retrieved November 30, 2021, from https://www.cde.state.co.us/nutrition/homedeliveryforsfspsponsors.
5. Meltzer, E., & Fu, A. (2021, June 29). Colorado schools are getting more money. bigger changes could be on the way. Chalkbeat Colorado. Retrieved November 30, 2021, from https://co.chalkbeat.org/2021/6/29/22549459/colorado-school-funding-changes-analysis.
6. School meal trends & stats. (n.d.). Retrieved November 30, 2021, from https://schoolnutrition.org/aboutschoolmeals/schoolmealtrendsstats/.
7. Strategies and challenges in feeding out-of-school students. (n.d.). Retrieved November 30, 2021, from https://www.urban.org/sites/default/files/publication/102095/strategies-and-challenges-in-feeding-out-of-school-students.pdf.
8. Students qualifying for free or reduced price lunch: Kids count data center. KIDS COUNT data center: A project of the Annie E. Casey Foundation. (n.d.). Retrieved November 30, 2021, from https://datacenter.kidscount.org/data/tables/469-students-

qualifying-for-free-or-reduced-price-lunch?loc=7&loct=10#detailed/10/1278-1457/false/574,1729,37,871,870,573,869,36,868,867/109,110,111/11515,7665.

9. Nate Frentz and Zoë Neuberger (2012, September 6). Key steps to improve access to free and reduced-price school meals. Center on Budget and Policy Priorities. Retrieved November 30, 2021, from https://www.cbpp.org/research/key-steps-to-improve-access-to-free-and-reduced-price-school-meals.