Problem Statement / Question of Interest

My project aims to find patterns between school assessment scores and noise pollution in Colorado state. Upon beginning my project I discovered there are very few articles on the distance noise pollution travels. The articles about noise pollution explain that there are too many variables to track the noise pollution because every location is unique, being affected by noise mitigation devices, trees, and even weather. For these reasons most studies about noise pollution focus on the effects of noise on health using specific subjects like students, animals in a certain area, sleep patterns in humans, etc. Those studies and experiments observe which levels are reasonable for humans and animals, and which levels have negative health impacts. My project compares schools near noise pollution sources to schools not near noise pollution, while controlling for income.

Data sources & description

This project uses shapefile data and tabular data.

I retrieved median household income data for each school district from American Community Survey. http://proximityone.com/sd_co.htm

I retrieved a shapefile for major roadways from the Bureau of Transportation Statistics. https://data-usdot.opendata.arcgis.com/datasets/national-highway-planning-network?geometry=124.587%2C29.226%2C2.067%2C67.338

I retrieved airport locations and school locations from the Colorado Department of Public Health and Environment. https://data-cdphe.opendata.arcgis.com/datasets/fec1a4755e7f454389dcd18e183c8e08_0

I used school scores from the Colorado Department of Education website. Excel file: CMAS: 2019 CMAS Math and ELA District and School Overall Results https://www.cde.state.co.us/assessment/cmas-dataandresults-2019

Data Conditioning and Transformation

In my project I use two sources of noise pollution: airports (including military air bases) and major roadways. Noises from major roadways are often mitigated by walls and trees, but unmitigated noise from a highway can still be above 50 decibels at 400 feet (City of San Diego study: Noise Element). I used a buffer zone of 350 feet on major roadways. In studies about airport noise pollution, the 50 decibels mark generally travels 3.5 to 5.5 kilometers depending on the environment (Faiyetole and Toyin, Predicting the aircraft). I used a buffer zone of 3.5 kilometers for the airports.

The Colorado Department of Education 2019 CMAS spreadsheet shows each school's scores broken out by grade as well as the mean score for the entire school per subject. For my project I exclusively used the mean score of the math results for the entire school.

I scrubbed the tabular data with the median family incomes for school districts and matched the schools to their school districts, made it suitable to upload as a table to arcmap, and then joined that information with the school locations shapefile. I started with 2,325 school locations, I eliminated all schools that didn’t have matching school information or test scores leaving me with 1,313 schools. I created a noise pollution buffer zone layer by merging the buffers built on the roadway and airport shapefiles. Using the select by location tool, I exported the 221 schools that fell within the noise pollution buffer zones, and also exported the 1,092 schools outside of the noise.

Analysis Method

I exported the 2 groups of schools (noise pollution and no noise pollution) into separate excel files and divided the schools up by income groups in $20,000 brackets. ($20,000 - $40,000, $40,000 - $60,000, etc). I built a linear regression model using the statsmodels library. In model 1, I predicted test scores using median income and whether or not the school is in the noise buffer (as a dummy variable). In model 2, I added an interaction term (median income*inside buffer). I also used a Welch's t-test to find the difference in means.

Analysis Results

The result of the Welch's t-test indicates that above a certain income threshold the noise pollution may have an effect on test scores. Essentially this says that students that attend a school with a lower income, that the poverty or lack of investment per student is a more important indicator of their test scores until the $60,000 bracket. Above $60,000 other factors, potentially noise pollution, have an effect on the test scores.

in buffer	20k-40k	40k-60k	60k-80k	80k-100k	100k+
mean	722	730	734	738	733
standard dev	11.05271	14.56761	14.599647	18.46761	13.75682
N	16	106	71	19	9
s	2.763179	1.414931	1.7326594	4.23676	4.585605
s2	7.635156	2.002031	3.0021084	17.95014	21.02778
v	15	105	70	18	8
1/v	0.066667	0.009524	0.0142857	0.055556	0.125
s4	58.29561	4.008128	9.012655	322.2075	442.1674

out buffer	20k-40k	40k-60k	60k-80k	80k-100k	100k+
mean	726	729	739	743	747
standard dev	12.28149	14.32432	15.05606	12.7222	10.41291
N	75	412	426	107	72
s	1.418144	0.705708	0.729468	1.229901	1.227174
s2	2.011133	0.498024	0.532124	1.512657	1.505955
v	74	411	425	106	71
1/v	0.013514	0.002433	0.002353	0.009434	0.014085
s4	4.044655	0.248028	0.283156	2.288133	2.267902

	20k-40k	40k-60k	60k-80k	80k-100k	100k+
std dev of two samples	3.105848	1.581156	1.8799555	4.411666	4.746971
s4	93.05089	6.250276	12.4908	378.8004	507.7691
1/v	0.042354	0.006204	0.0103611	0.047313	0.108913
v	23.6108	161.1888	96.514823	21.13606	9.181605
v rounded	23	161	96	21	9
t	-1.28789	0.632449	-2.6596375	-1.13336	-2.94925
p value ranges	0.2 - 0.3	0.5-0.6	0.005 - 0.01	0.2 - 0.3	0.01 - 0.02

P Values for the two-tailed test

In the linear regression, the results of Model 1 indicate that income raises test scores and being in the noise buffer zone lowers it, and these results are statistically significant. Model 2 indicates that income raises test scores but it raises less for the schools in the noise buffer zone, and the buffer zone is not statistically significant at the 95% level.

Discussion & Conclusion

There are several changes I would make to repeat this project. I lost some schools when matching school district names with income because the median income school names did not match the school names on the shapefile. I did my best to match them myself, for example some schools changed names and some schools merged, so in the future I want better quality school income data. I also think it would be interesting to break out my results by grade levels, or just focus on highschool grades 9-11 rather than all grades. I want to make smaller income brackets, but I would need more schools, as some of my income brackets are already too small and therefore not statistically significant. Perhaps I could use a state with more schools or focus on an area with student level data, increasing my n.

I understand the nature of noise pollution makes this project's results questionable, and there are factors to consider when seeing these buffer zones. For example, it's possible that air pollution from roadways and airports is creating this pattern rather than noise pollution. It's also likely that some schools have better noise mitigation than others, whether inside or outside the noise buffer zone.

I struggle with making the maps visually appealing. Software like Arcmap allowed me to find the results and was crucial to this project, but once I pulled the data from arcmap it's as if my maps lost their story. Attempts to tell the story of my results with my maps ended up telling a

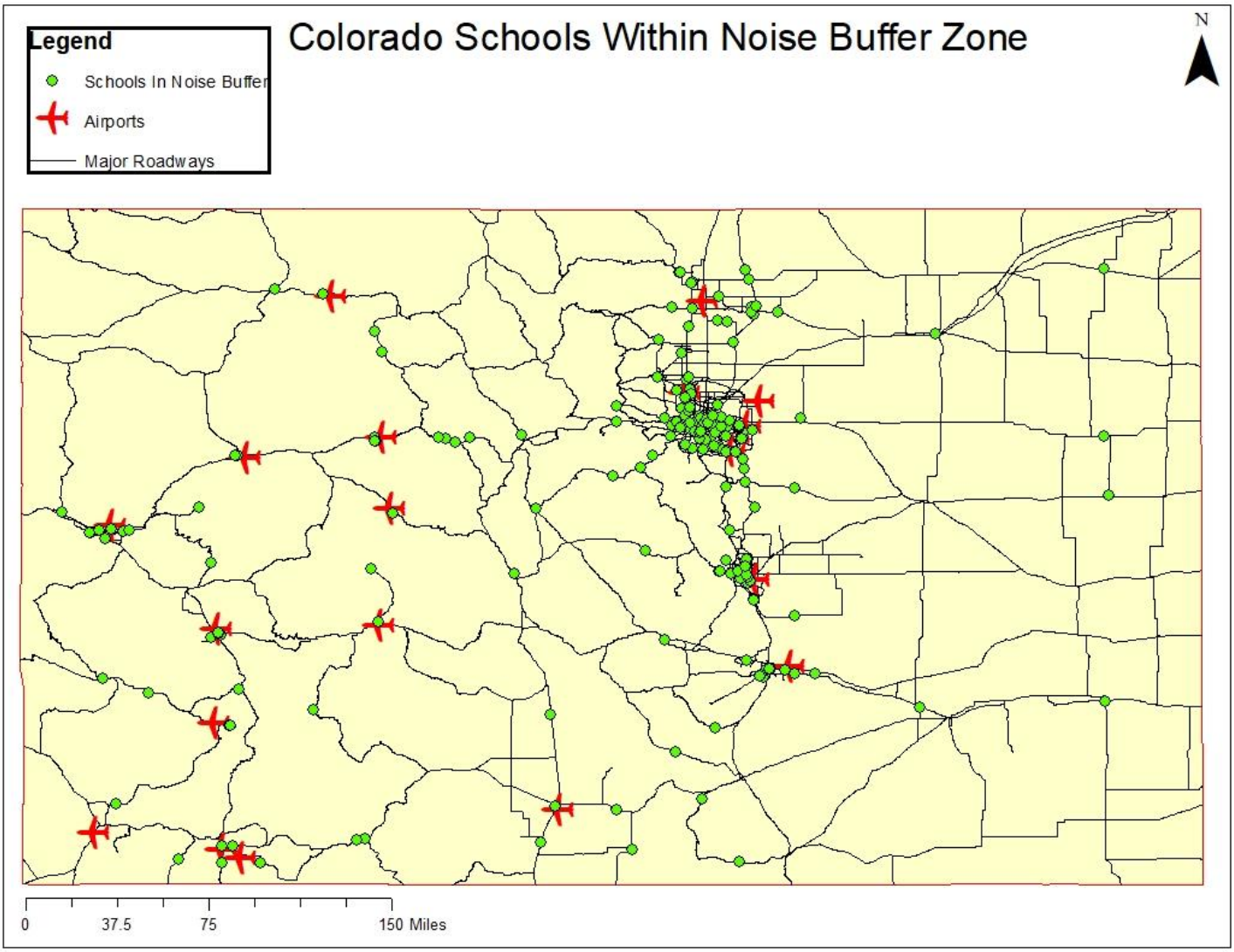
different story, for instance, I inadvertently ended up with a map of incomes within school districts and it didn't convey what my results found and instead indicated where wealth was - not my intention.

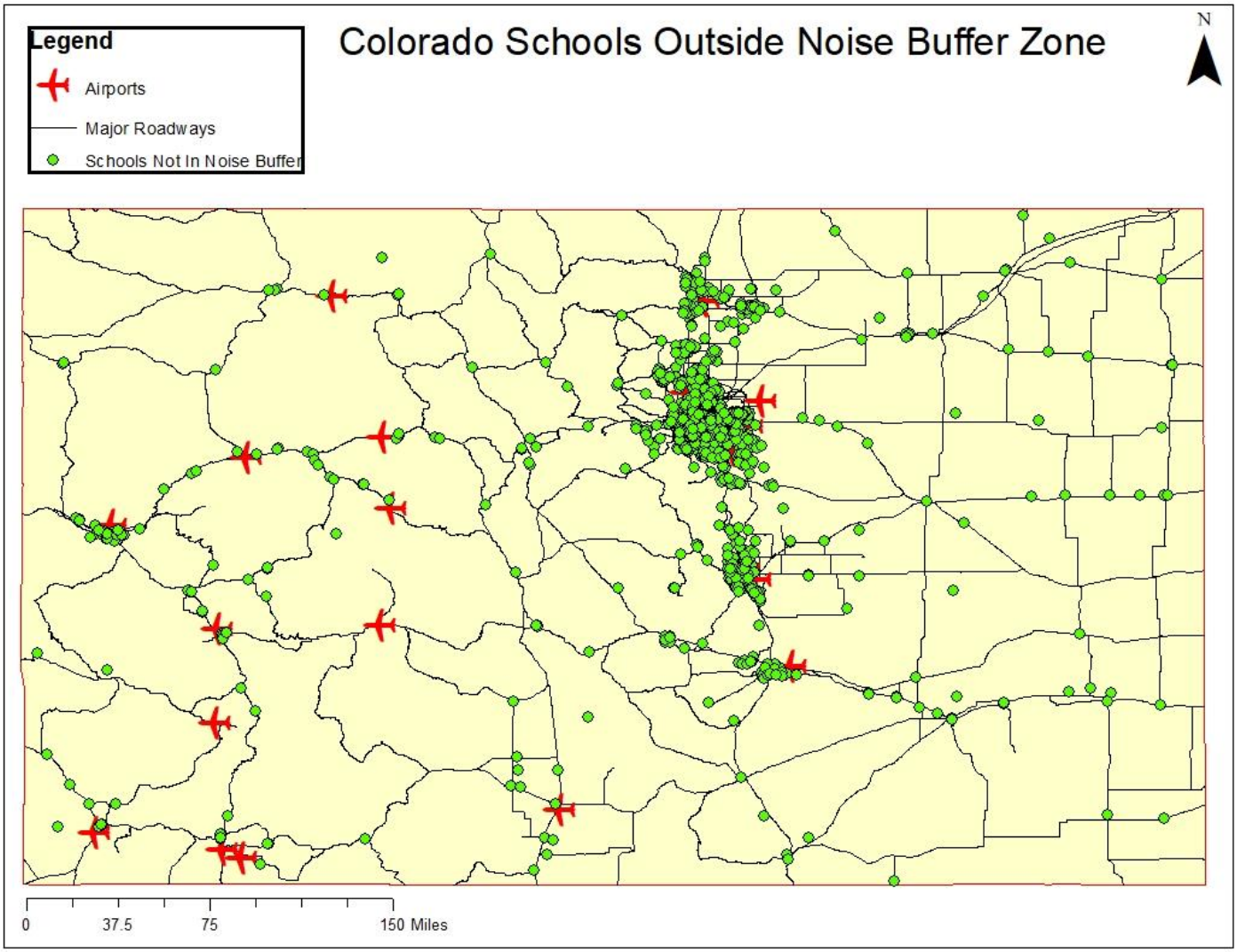
Overall, I'm impressed by the results of my project, and I wish to research noise pollution further so I may expand and change this project. It is clear that there are patterns in my data and that it requires a more comprehensive investigation.

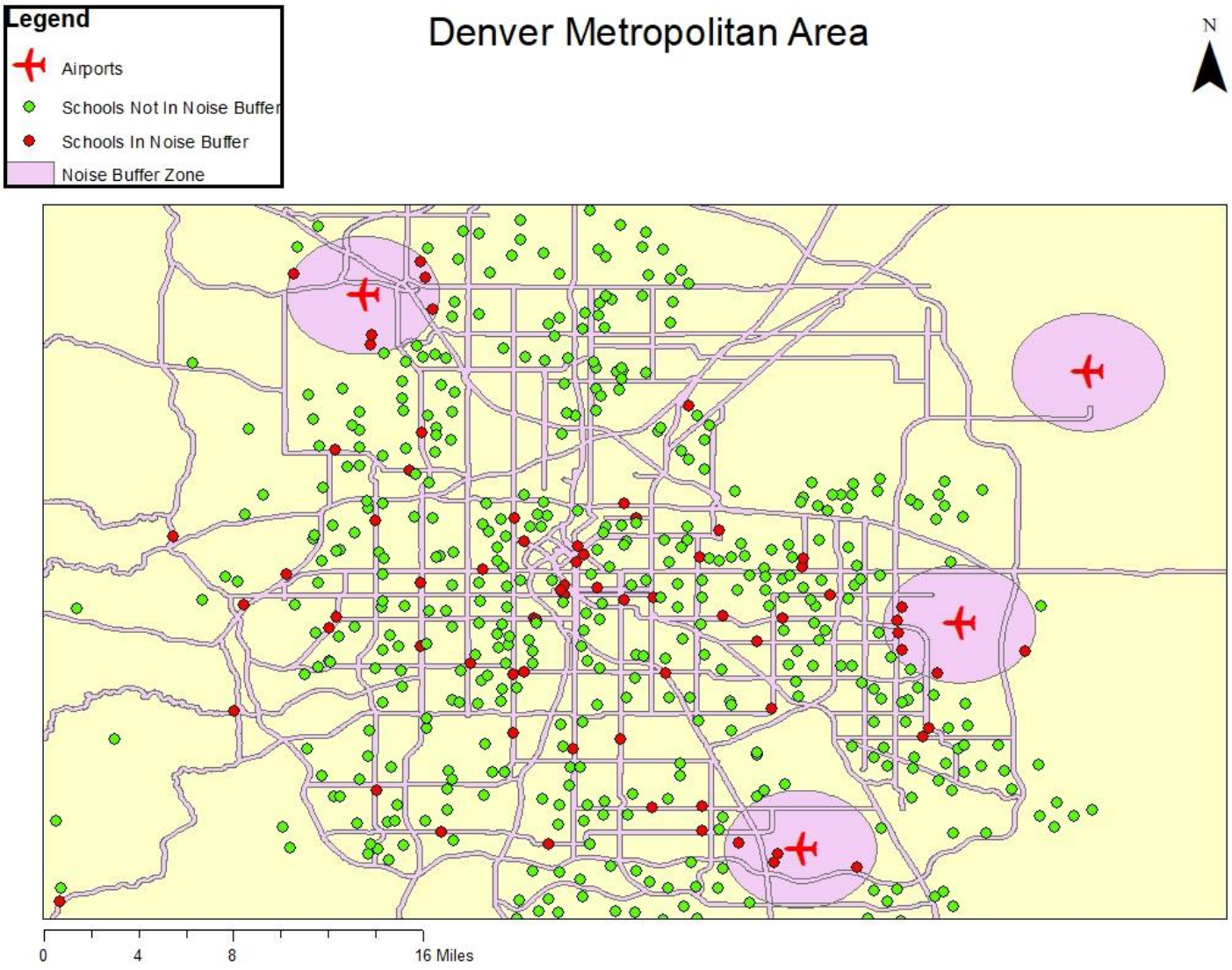
References & Citations

1. Jarup, Lars, et al. "Hypertension and exposure to noise near airports (HYENA): study design and noise exposure assessment." Environmental Health Perspectives, vol. 113, no. 11, 2005, p. 1473+. Gale Academic OneFile, link.gale.com/apps/doc/A140014426/AONE?u=cuny_baruch&sid=AONE&xid=7bc3d0f
2. Ayodele Adekunle Faiyetole, Ibrahim Taofiq Toyin. "Predicting the aircraft take-off noise level" Department of Transport Management Technology, Federal University of Technology Akure. https://www.jvejournals.com/article/20588/pdf
3. FAA.GOV https://www.faa.gov/regulations_policies/policy_guidance/noise/basics/
4. City of San Diego General Plan "Noise Element" https://www.sandiego.gov/sites/default/files/ne_2015.pdf
5. Jeffrey P. Cohen, Cletus C. Coughlin. "SPATIAL HEDONIC MODELS OF AIRPORT NOISE, PROXIMITY, AND HOUSING PRICES". JOURNAL OF REGIONAL SCIENCE, VOL. 48, NO. 5, 2008, pp. 859–878

Jacob Robert van Winkle
Noise Pollution and Colorado School Scores


Legend
Schools In Noise Buffer
Airports
Major Roadways
Colorado Schools Within Noise Buffer Zone
N
0
37.5
75
150 Miles


Legend
Airports
Major Roadways
Schools Not In Noise Buffer
Colorado Schools Outside Noise Buffer Zone
N
0
37.5
75
150 Miles


Legend
Airports
Schools Not In Noise Buffer
Schools In Noise Buffer
Noise Buffer Zone
Denver Metropolitan Area
N
0
4
8
16 Miles