Proceedings of the 7th International Conference on Statistics: Theory and Applications (ICSTA'25)
Paris, France – August 17 -19, 2025
Paper No. 101
DOI: 10.11159/icsta25.101

Pre, Peri, and Post-Covid Statistical Analysis of Restaurant Inspection Scores: Preliminary Insights

Nicholas V. Scott[1], Sarah Jensen[2], and Tania Nur[2]
[1]Apogee Engineering, LLC, Science and Technology Group
2611 Commons Blvd., Beavercreek, OH, 45324, USA
nscott.gso@gmail.com; sarahjensen@franklincountyohio.gov; tanianur@franklincountyohio.gov
[2]Franklin County Public Health, Environmental Health
280 East Broad Street, Columbus, Ohio, 43215, USA

Abstract - Restaurant inspection score ratings serve as a vital evaluation tool for adherence to food safety codes. This analysis examines inspection scores across 40 zip codes in Austin, Texas during pre-Covid (March 2014 to March 2017) and peri-post-Covid (August 2021 to August 2024) periods. The data reveals spatio-temporal variability in restaurant compliance with health regulations, offering insights for future health crises. During the pre-Covid period, the mean inspection score was higher by 5 points than in the peri-post-Covid period, indicating better adherence to pre-pandemic food safety regulations. The percentage of zip codes with inspection scores of 70 and below rose from approximately 50% to 78% over the 2 time periods, underscoring the pandemic's widespread impact on restaurant performance and providing a useful statistic for future planning. In addition, the coefficient of variance increased from [0.04 0.08] to [0.1 0.2], reflecting more frequent health code infractions which are most likely due to resource shortages for food distribution businesses in the peri-post-Covid timeframe. Principal component analysis indicated a reduction in dimensional time scale for pre-Covid inspection score performance, signifying healthy restaurant behavior characterized by time scale redundancy. However, no such reduction was observed in the peri-post-Covid period, suggesting that normal operational dynamics were disrupted. The analysis also revealed a consistent five-dimensional structure for inspection scores across both timeframes, indicating that the same underlying restaurant institution framework influenced spatial variability in inspection scores. This preliminary analysis highlights the sensitivity of inspection scores to external disruptions caused by the Covid-19 pandemic and underscores the need for further data to establish more definitive conclusions.

***Keywords*:** Covid-19, food safety, inspection scores, principal component analysis, restaurants, skewness, zip codes, variance

1. Introduction

The United States Food and Drug Administration (FDA) is the federal entity responsible for the explicit delineation of food codes whose adherence by food distribution establishments helps to prevent food borne illness and massive food disease outbreaks ensuring general public health [1, 2]. Even with the enforcement of food health codes under normal circumstances, recurrent food health code violations remain [3]. The recent Covid-19 pandemic was a worldwide health crisis affecting the dynamics of many industries and organizations including those associated with the production and distribution of food across small and large spatial areas. One industry within the food distribution genre that was significantly affected was the restaurant business industry where an untold number of restaurants were forced to change the methods by which they operated by government mandated law. This included not only how patrons of restaurants could sit within food distribution establishments but also how food could be served and when by businesses. Of particular importance to health officials who specialize in the monitoring of safe food distribution practices of restaurants is understanding how restaurant health violation distributions changed during the Covid-19 period. Measurable changes in the dynamics of restaurants before and after the pandemic is an area of great interest to health officials working at the state and local governmental levels, since such information is crucial for the facilitation of policy creation for effectively dealing with future health crises. However, attaining pervasive information regarding restaurant food health practices, especially during this period of great health adversity, has been difficult. This has been due to many reasons and factors including laws regarding what can legally be measured within a restaurant regarding food health code adherence, the lack of desire by owners to freely share restaurant health practices, and the closing of many restaurants unable to economically cope with stresses brought on by the crisis. One viable metric for the assessment of restaurant food safety practices is the restaurant inspection score whose analysis is the focus of this paper.

Even with the existence of the aforementioned barriers, different cities and counties have been willing to share their restaurant inspection score data allowing the public the opportunity to examine its structure for global assessment of food distribution health practice. The health department of the city of Austin, Texas is one such organization whose food distribution inspection data is analyzed here. The purpose of this brief paper is to characterize and explain the statistical trends and differences in publicly available inspection scores before, during, and after the Covid-19 health crisis. From the limited amount of data, the aim is to display and explain the statistical characteristics of the data using basic statistical science concepts and techniques. Though no trends explicated within this paper are to be taken as immutable truth as to the dynamics surrounding restaurant inspection scores, it is hoped that this work provides insight into understanding how inspection scores behave as a spatial random variable during a period of extreme stress [4, 5]. In addition, it is believed that the results demonstrate how the inspection score grading system can be seen as a useful way for identifying low inspection score areas which could be susceptible to food disease outbreaks [6, 7]. It is through understanding of the statistical behavior of inspection scores that insight can be attained and used to facilitate modulation of future restaurant health policy if a similar health crisis in the future occurs.

2. Data Structure

The data used in this statistical analysis was obtained from the official city of Austin, Texas data portal website responsible for the posting of open-source data for public use [8]. The data consisted of Excel formatted food establishment inspection scores accrued from restaurants and a wide variety of food distribution venues including delicatessens, school cafeterias, convenience stores, grocery stores, retail markets, gas stations, motels, bakeries, and bars situated in Austin, Texas. The Austin/Travis County Health and Human Services Department performs inspections of more than 4,000 food establishments in Austin, Texas which are required to be inspected twice a year. Businesses are evaluated on a scale of 0 to 100, where higher values denote greater adherence to the law, with scores below 70 being required to elevate their score by adherence to health inspection guidelines provided by the Texas Food Establishment Rules. Two different Excel spread sheet data sets for the pre-Covid period, and peri and post-Covid period consisted of columns designating food establishment name, zip code, address location, inspection time, and inspection scores. Each row was a data observation taken at different times. This data was pre-processed to produce m x n matrices where the m rows designate numerically ordered zip codes and the n columns designate consecutive 6-month intervals. The m and n dimensions were m=54 and n=6 for the pre-Covid time period and m=50 and n=6 for the peri and post-Covid time period. The pre Covid time period data set extended from fall 2018 to spring 2021. The peri and post Covid time period data set extended from fall 2021 to spring 2024. The number of zip codes shared during both Covid time periods was on the order of 95%.

For each data matrix associated with the different Covid time periods, three different feature matrices were created where the zip code and 6-month time intervals were arranged in numerically ascending order. The first feature matrix consisted of mean inspection score values for each 6-month time interval coordinate and zip code coordinate. The second and third feature matrices consisted of minimum inspection scores and coefficient of variation values respectively for each 6-month time interval coordinate and zip code coordinate. It is from these matrices that statistical calculations for the two Covid time periods were performed to characterize inspection score structure with the purpose of eliciting differences and similarities. It is again noted that the analyzed inspection score information included not only restaurant inspection scores but scores from all types of food distribution establishments. However, the predominant food serving entity in the inspection score data were restaurants and it is their data dominance that provides credence to the statistical results and trends provided here.

3. Statistical Analysis Results:

Mean inspection score values over the 6-month time intervals for all of the zip codes were computed for both Covid time periods and displayed using box plots. The zip codes used in the analysis for each Covid time period were approximately the same with 48 zip codes shared during both periods. Figure 1a)-b) displays boxplots for the pre-Covid time period, and peri and post-Covid time periods respectively. The figures show mean inspection score distributions where the pre-Covid time period has an average mean inspection score value which is higher by 5 inspection score points than the average mean inspection score value for the peri and post-Covid time period. This suggests that restaurants on average had been adhering to food safety regulations much better during the pre-Covid time period. It is noted that factors such as inspection score grade inflation do exist which may have caused higher than usual pre-Covid scores [9]. This is due to the impact of repetitive interaction of food sanitarians with restaurants which can introduce a bias where inspections scores are not as low as they should be.

Examination of the minimum inspection score statistics provides additional insight. Figure 1c)-d) are boxplots of the minimum inspection score values across zip codes showing larger amounts of low inspection scores during and after the Covid pandemic than before. The number of zip codes having inspection scores of 70 and below before the pandemic is approximately 50%. This number increases to 78% for the peri and post-pandemic period. The uniform drop in inspection score values across all zip codes demonstrates not only how pervasive the effect of the pandemic across all zip codes was but also provides an expected percent inspection score drop due to the health crisis. The percent drop is a practical and useful statistic establishing a Bayesian prior for future planning concerning similar health situations.

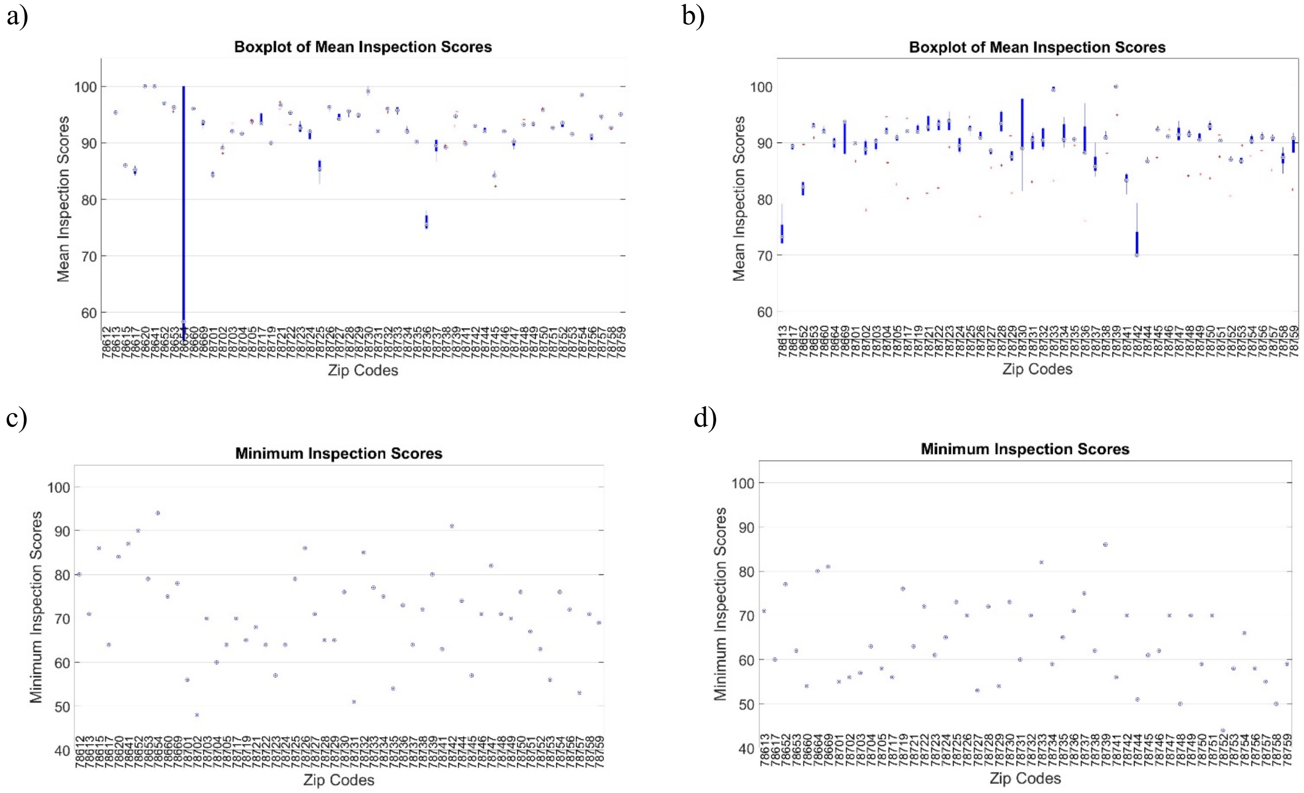

Figure 1: Boxplot of mean inspection score values across zip codes for a) pre-Covid time period and b) peri and post-Covid time period. Boxplot of minimum inspection score values for c) pre-Covid time period and d) peri and post-Covid time period. Over 50 zip codes appear on x-axis. Central blue dot indicates mean or mean minimum values for each zip code. Bottom and top edges of the blue thick line indicate the 25th and 75th percentiles respectively for the mean inspection score distributions. Whiskers extend to the most extreme data points not considered outliers.

Many inspection score outliers exist during the second Covid time period which are indicative of the extreme variability caused by the Covid pandemic shut down. It is strongly conjectured that the shutdown was responsible for many restaurants inability to adhere to health code regulations due to extreme stress placed on internal systems needed for restaurants to function properly. Figure 2a)-b) are boxplots of the coefficient of variation for inspection score values across zip codes. The pre-Covid time period possessed a variance range of [0.04 0.08]. The variance range increased to [0.1 0.2] for the peri and post-Covid time period, where the large increase again reflects health code infractions committed by restaurants influenced by factors cited above.

The large degree of variance is also reflected in the skewness distribution across zip codes. Figure 2c)-d) displays the skewness value distribution where the frequency of inspection score skewness values were tabulated over the entire zip code region for both Covid time periods. The pre-Covid time period displays a skewness peak at –0.5 whereas the peri and post-Covid period shows a peak at -1.7. The decrease in the mean value for the frequency of skewness value distributions over the two time periods is directly attributed to the increase in extreme variability of inspection scores. It is worthy of note that increases in food health violations contributing to the skewness are caused by factors which are not always due to restaurant ineptitude. The inability of the food distribution businesses to obtain the resources needed to perform their function within the parameters of the health code laws was a serious reason for lower inspection scores and variability during the latter Covid time period.

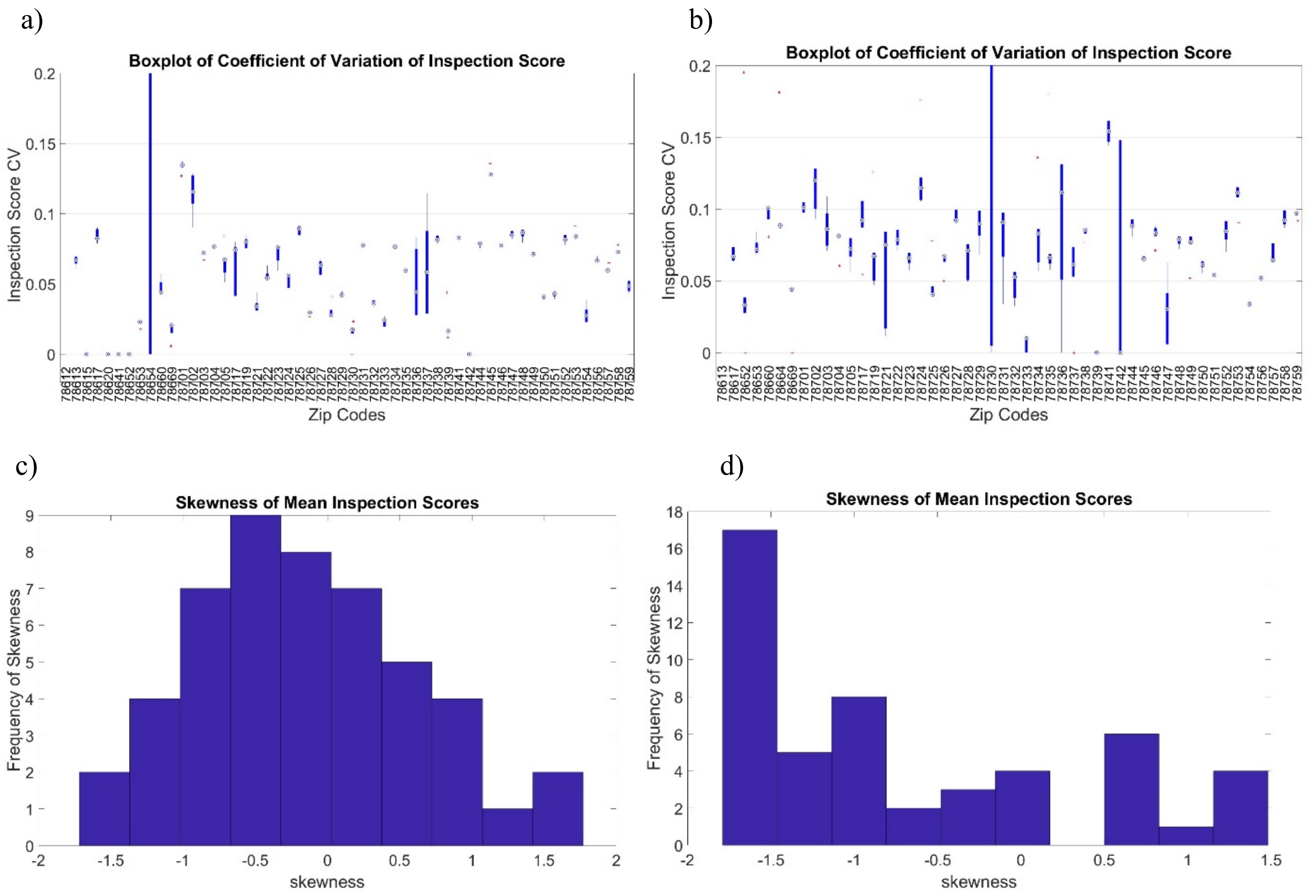

Figure 2: Boxplot for coefficient of variation for inspection scores across zip codes for a) pre-Covid time period and b) peri and post-Covid time period. Blue circles designate mean coefficient of variance values for inspection scores for each zip code. Over 50 zip codes appear on x-axis. Central mark indicates the median, and the bottom and top edges of the blue thick line indicates the 25th and 75th percentiles respectively. Whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the red '+' marker symbol. Skewness histogram for inspection score values for c) pre-Covid time period and d) peri and post-Covid time period. Frequency values on y-axis and skewness values on x-axis.

The 6 X 6 temporal covariance matrix was computed using mean inspection score matrices for both Covid time periods [10]. This is shown in Figure 3a)-b). The pre-Covid time period covariance matrix shown in Figure 3a) shows many time correlations. Among these are a correlation between the second 6-month time interval and the sixth 6-month time interval. The third and the fifth 6-month time intervals are also highly correlated. These correlations can be explained. The first correlation represents a spring seasonal time-scale correlation while the second represents a fall time-scale correlation. The peri and post-Covid time period covariance matrix shown in Figure 3b) shows strong correlations for the first to the fourth 6-month time intervals. This correlation does not reflect true restaurant health code violation dynamics across zip codes but rather is indicative of how restaurant operations were stalled during the Covid-19 shutdown. The consistently high correlation shown across the first four 6-month time intervals is indicative of the influence of the health pandemic and how it masks any restaurant inspection score variability.

Principal component analysis was performed on the temporal 6 X 6 dimension mean inspection score covariance matrices for both Covid time periods. Principal component analysis provides an eigenvalue spectrum delineating the energy distribution as a function of temporal dimensional scale [11,12]. Figure 3c) shown the eigenvalue spectrum for the covariance matrix associated with the pre-Covid time period with a noticeable elbow in the spectrum at the fourth spectral dimension. The eigenvalue spectrum suggests that only four 6-month temporal dimensions are necessary to describe the variance in the pre-Covid inspection scores. The dimensional collapse from the full 6 temporal dimensions to 4 temporal dimensions reflects how the full 6 temporal dimensions are not necessary to explain the inspection score variance. If restaurants are adhering to health codes, some dimensional reduction and redundancy is expected due to seasonal variability in health code violations, a fact known to sanitarians responsible for inspecting food distribution businesses. Figure 3d) shows no noticeable elbow suggesting that all temporal dimensions are needed to explain the inspection score variability. This is not surprising since inspection score noise pervades the complete three-year time span during this period. Such noise disallows a dimensional collapse where inspection scores occupying extreme value regions as well as values near the mean value exist over a 3-year data time span.

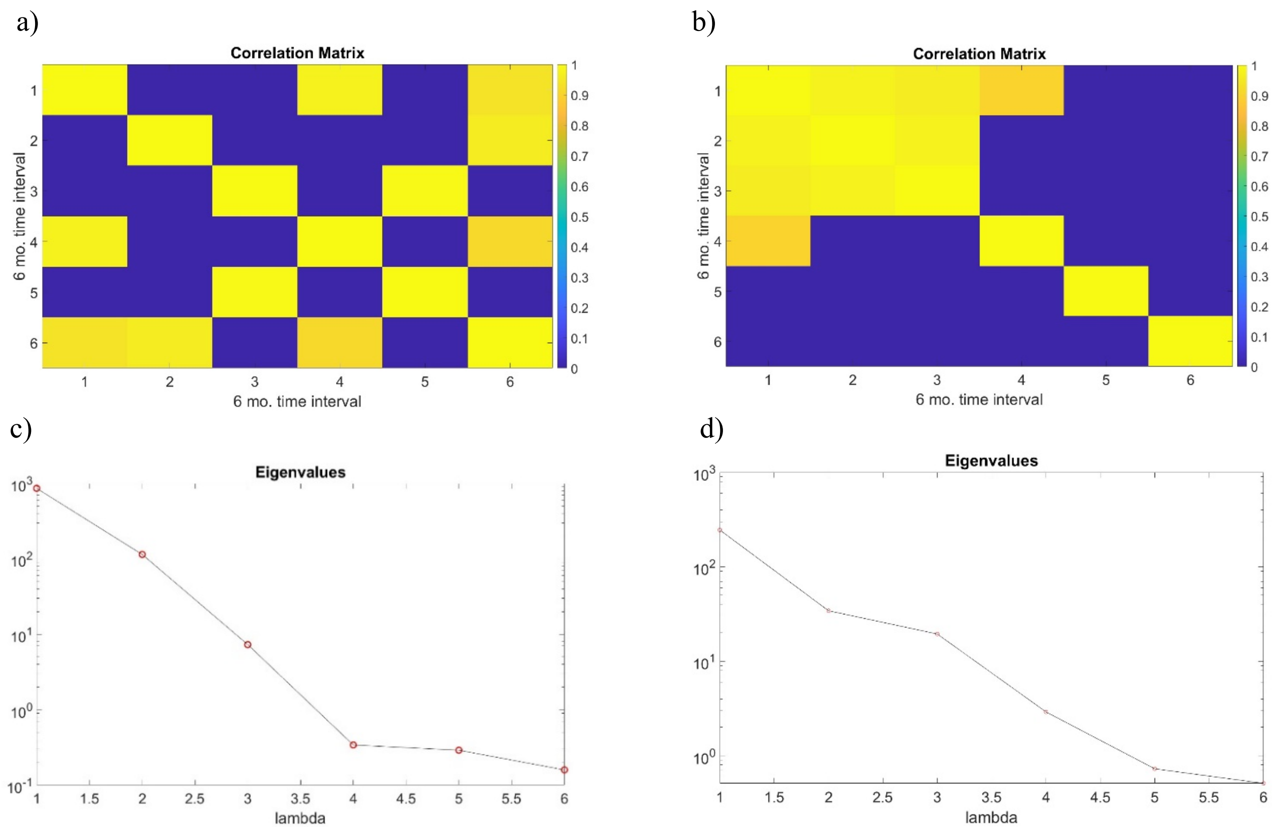

Figure 3: Temporal covariance matrices for the a) pre-Covid and b) peri and post-Covid time periods. Temporal eigenvalue spectra for the c) pre-Covid and d) peri and post-Covid time periods. Six temporal 6-month time intervals denoted by lambda on the x-axis span 3 years starting from spring 2017 and spring 2021 respectively. The y-axis denotes spectral power in arbitrary units.

The spatial zip code covariance matrix was also computed using mean inspection score matrices for both Covid time periods. This is shown in Figure 4 a)-b). The pre-Covid time period spatial covariance matrix shown in Figure 4a) shows that 17 pairs of zip codes share substantial covariance over the complete 3-year time period. Approximately 90% of the 17 zip code pairs were not in close proximity. The significant covariance could be due to noise but may also reflect real covariant behavior. Such covariance may be due to many factors including the same inspectors assigned to specific zip code regions. In addition, areas with strong covariance may also reflect regions which share the same food distribution types where health code infractions have a tendency of being similar. No definitive covariance pattern is evident disallowing extensive inference from the covariance matrix values. The covariance matrix for the peri and post-Covid time period possesses a speckle pattern clearly suggesting large amounts of spurious correlations between zip codes. This covariance matrix reflects the large amounts of statistical noise existing within the inspection score data which obscures true restaurant and food establishment health code violation dynamics across zip codes during the Covid-19 shutdown.

Principal component analysis was performed on the spatial covariance matrix for both Covid time periods. The spatial covariance matrices for the pre-Covid, and peri and post-Covid time periods possessed m x m dimensions of m=54 and m=50 respectively. The eigenvalue spectra, delineating the energy distribution as a function of dimensional scale, are shown in Figure 4c-d) for the two Covid time periods. Both eigenvalue spectra depict a noticeable elbow in their spectra at the transition between the 5th and 6th zip code spatial dimensions with a more significant elbow drop experienced for the peri and post-Covid time period. Though previous analysis of the temporal inspection score dynamics suggest noise that is spread out over time, analysis of the spatial inspection score dynamics suggests noise that is more segregated with respect to spatial scale. In other words, the principal component analysis results suggests that even with the injection of high amounts of inspection score noise, from a linear perspective five virtual zip codes are responsible for carrying large amounts of inspection score variance for both time periods. This result suggests that the spatial distribution of businesses and how they respond to health code inspections in the Austin, Texas is possibly robust across both Covid time periods. This principal component analysis-based spatial scale could possibly serve as an inspection score latent spatial label parameterizing and uniquely characterizing the Austin, Texas area.

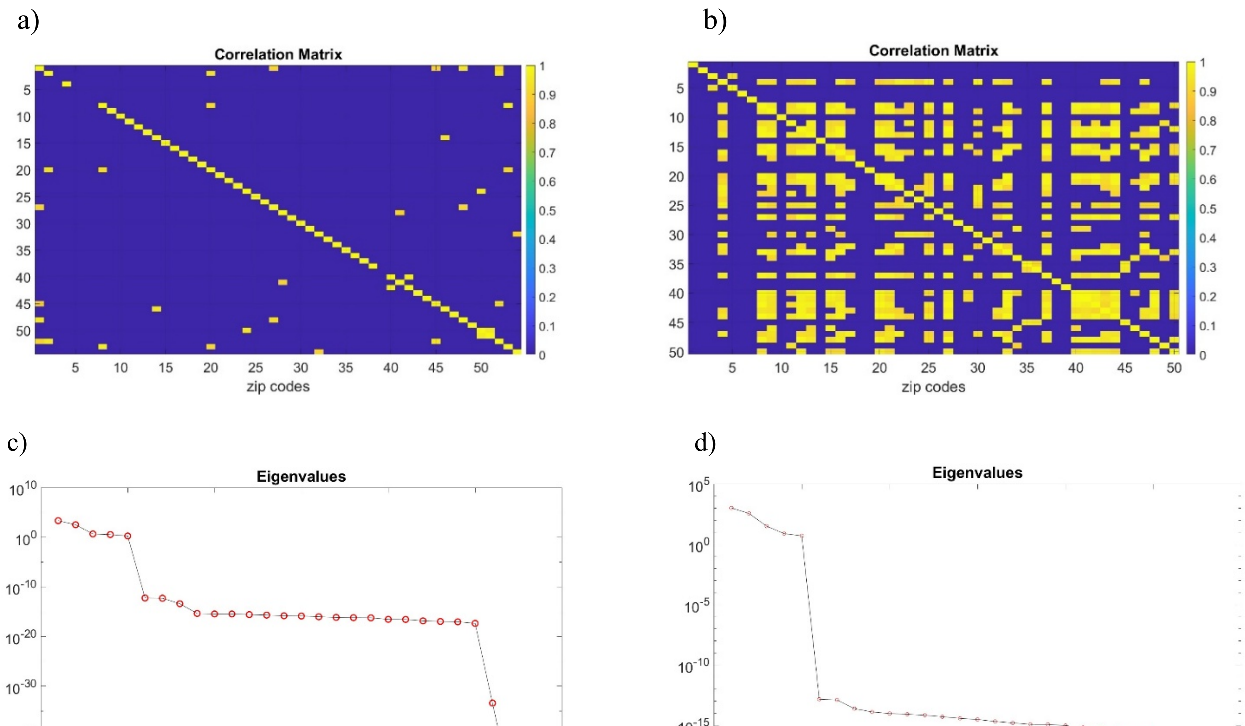

Figure 4: Spatial zip code covariance matrices for the a) pre-Covid and b) peri and post-Covid time periods. Spatial eigenvalue spectra for the c) pre-Covid and d) peri and post-Covid time periods. Number of zip codes for each Covid time period is 54 and 50 respectively. The zip code regions are denoted by lambda on the x-axis. The y-axis denotes spectral power in arbitrary units. The percentage of shared zip codes is 95%.

4. Conclusions:

Many health science experts believe that future health pandemics are highly likely and that prudent preparatory measures are the best bulwark for addressing how restaurants should deal with the situation. Part of the process for future preparation is the development of robust food business health code policy which rests on rigorous understanding of how restaurants and other food distribution businesses behave with respect to adherence to health code policy. This work provides insight into how to explore regional inspection score variability, providing metrics for understanding how different regions perform with respect to health code adherence under extreme stress provided by a health pandemic. Minimum inspection score distributions across zip codes are shown as numerical values in Figure 5 a)-b) for the pre-Covid, and the peri and post-Covid time periods respectively. The minimum inspection score trends across zip code regions for both the pre, and peri and post-Covid time periods possess large amounts of spatial variability with a noticeable number of low values in the central region of Austin, Texas. This is not surprising since this is the area associated with many food distribution businesses. With such a high density of food businesses the probability of low values increases significantly. There are areas that experience increases in minimum inspection score values from the pre-Covid to the peri and post-Covid time periods such as zip codes 78719, 78735, 78733, and 78739. Overall, however, there is a slight decrease in the minimum inspection score value for this transition period highlighting a global spatio-temporal trend worthy of note.

a) b)

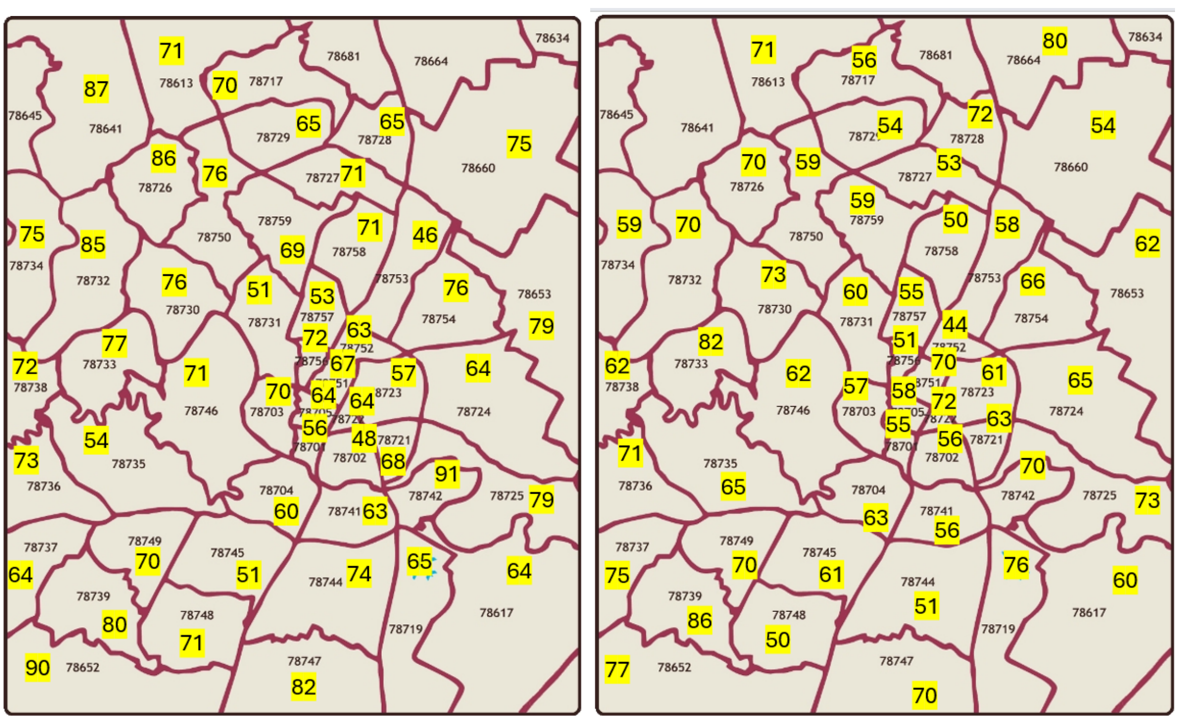

Figure 5: Spatial map of minimum inspection scores across Austin, Texas zip codes for the a) pre-Covid time period and b) peri and post-Covid time periods. Zip codes labeled. Minimum inspection score values displayed in yellow boxes along with labeled zip codes.

It is conjectured that the observed spatial variability could be smoothed producing more homogeneity via improved education, so businesses are better equipped to deal with future health situations as it pertains to food distribution. In fact, in many states there has been an increased push to provide food handling and safety management to restaurant managers under the hypothesis that such education will function to lower food health infractions [13, 14, 15]. It is the coupling of the statistical trends observed and learned here with such education that bodes well for weathering future situations to ensure public health at the county, state, and national levels.

References

[1] D. Dewey-Mattia, K. Manikonda, A. J. Hall, M. E. Wise, and S. J. Crowe, “Surveillance for Foodborne Disease Outbreaks-United States, 2009-2015,” *Center for Disease Control and Prevention, MMWR Surveillance Summaries*, vol. 67, no.10, pp. 1-11, 2018.
[2] Adoption of the FDA Food Code by State and Territorial Agencies Responsible for the Oversight of Restaurants and/or Retail Food Stores, 2023.
[3] M. L. Phillips, B. L. Elledge, H. G. Basara, R. A. Lynch, and D. T. Boatright, “Recurrent critical violations of the food code in retail food service establishments,” *Journal of Environmental Health*, vol. 68, no. 55, pp. 24–30, 2006.
[4] Y. Lee, L. Pennington-Gray, and J. Kim, “Does location matter? Exploring the spatial patterns of food safety in a tourism destination,” *Tourism Management*, vol. 71, pp. 18–33, 2019.
[5] D. L. T. Ng, M. A. Harris, T. H. Oiamo, and I. Young, “Spatial Distribution and Characteristics of Restaurant Inspection Results in Toronto, Ontario 2017–2018,” *Food Protection Trends,* vol. 40, pp. 232–240, 2020.
[6] R. L. Petran, B. W. White, and C.W. Hedberg, “Health Department Inspection Criteria More Likely To Be Associated with Outbreak Restaurants in Minnesota,” *Journal of Food Protection*, vol. 75, pp. 2007–2015, 2012.
[7] R. L. Petran, B. W. White, and C.W. Hedberg, “Using a Theoretical Predictive Tool for the Analysis of Recent Health Department Inspections at Outbreak Restaurants and Relation of This Information to Foodborne Illness Likelihood,” *Journal of Food Protection* ,75, pp. 2016–2027, 2012.
[8] “Official City of Austin data portal,” https://data.austintexas.gov/Health-and-Community-Services/Food-Establishment-Inspection-Scores/ecmv-9xxi/about data, October 10, 2024.
[9] B. Kovács, D. W. Lehman, and G. R. Carroll, “Grade inflation in restaurant hygiene inspections: Repeated interactions between inspectors and restaurateurs,” *Food Policy,* vol 97, no. 101960, DOI:10.1016/j.foodpol.2020.101960, 2020.
[10] A. Zagidullina, *High-Dimensional Covariance Matrix Estimation: An Introduction to Random Matrix Theory (SpringerBriefs in Applied Statistics and Econometrics)*. Cham, Switzerland: Springer, 2021.
[11] E. Alpaydin, *Introduction to Machine Learning, 2nd ed*. Boston, MA: MIT Press, 2010.
[12] W. L. Martinez, and A. R. Martinez, *Computational Statistics Handbook with MATLAB, Third Edition*. Boca Raton, FL: Chapman and Hall/CRC, 2015.
[13] A. Burke, M. R. Manes, L. Liu, and M. S. Dworkin, “Do Certified Food Manager Knowledge Gaps Predict Critical Violations and Inspection Scores Identified during Local Health Department Restaurant Inspections?,” *Food Protection Trends,* vol. 34, pp.101–102, 2014.
[14] J. Harris, “Certified Food Safety Manager Impact on Food Inspection Citations,” *Journal of Environmental Health*, vol. 80, pp. 16–21, 2017.
[15] S. C. Cates, M. K. Muth, S. A. Karns, M. A. Penne, C. N. Stone, J. E. Harrison, and V. J. Radke, “Certified kitchen managers: Do they improve restaurant inspection outcomes? *Journal of Food Protection,* vol. 72, no. 2, pp. 384–391, 2009.